



Cochrane
Library

Cochrane Database of Systematic Reviews

Imaging modalities for the non-invasive diagnosis of endometriosis (Review)

Nisenblat V, Bossuyt PMM, Farquhar C, Johnson N, Hull ML

Nisenblat V, Bossuyt PMM, Farquhar C, Johnson N, Hull ML.
Imaging modalities for the non-invasive diagnosis of endometriosis.
Cochrane Database of Systematic Reviews 2016, Issue 2. Art. No.: CD009591.
DOI: [10.1002/14651858.CD009591.pub2](https://doi.org/10.1002/14651858.CD009591.pub2).

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS	4
BACKGROUND	12
Figure 1.	16
OBJECTIVES	17
METHODS	17
RESULTS	22
Figure 2.	22
Figure 3.	24
Figure 4.	25
Figure 5.	28
Figure 6.	29
Figure 7.	30
Figure 8.	31
Figure 9.	32
Figure 10.	32
Figure 11.	33
Figure 12.	33
Figure 13.	34
Figure 14.	36
Figure 15.	36
Figure 16.	37
Figure 17.	37
Figure 18.	38
Figure 19.	39
Figure 20.	40
Figure 21.	41
Figure 22.	43
Figure 23.	44
Figure 24.	44
Figure 25.	44
Figure 26.	45
Figure 27.	46
Figure 28.	47
Figure 29.	48
Figure 30.	48
Figure 31.	49
Figure 32.	49
Figure 33.	50
Figure 34.	51
Figure 35.	52
Figure 36.	52
Figure 37.	53
Figure 38.	53
Figure 39.	54
Figure 40.	55
Figure 41.	56
Figure 42.	57
Figure 43.	57

Figure 44.	58
Figure 45.	58
Figure 46.	59
Figure 47.	60
Figure 48.	61
Figure 49.	62
Figure 50.	63
Figure 51.	65
Figure 52.	67
Figure 53.	68
Figure 54.	69
Figure 55.	70
Figure 56.	71
Figure 57.	72
Figure 58.	72
Figure 59.	73
Figure 60.	73
Figure 61.	73
Figure 62.	74
Figure 63.	74
Figure 64.	74
Figure 65.	75
Figure 66.	75
Figure 67.	76
Figure 68.	77
Figure 69.	78
Figure 70.	79
Figure 71.	79
DISCUSSION	79
AUTHORS' CONCLUSIONS	81
ACKNOWLEDGEMENTS	83
REFERENCES	84
CHARACTERISTICS OF STUDIES	96
DATA	232
Test 1. TVUS pelvic.	234
Test 2. TVUS ovarian.	234
Test 3. TVUS DIE.	234
Test 4. TVUS posterior DIE.	234
Test 5. TVUS* posterior DIE.	234
Test 6. TVUS USL.	234
Test 7. TVUS RVS.	234
Test 8. TVUS* RVS.	234
Test 9. TVUS vaginal.	234
Test 10. TVUS POD.	235
Test 11. TVUS anterior DIE.	235
Test 12. TVUS rectosigmoid.	235
Test 13. TVUS* rectosigmoid.	235
Test 14. TVUS bowel [ileum - rectum].	235
Test 15. TRUS ovarian.	235
Test 16. TRUS USL.	235
Test 17. TRUS RVS.	235
Test 18. TRUS vaginal.	235

Test 19. TRUS rectosigmoid.	235
Test 20. TRUS bowel [ileum - rectum].	235
Test 21. MRI pelvic.	235
Test 22. MRI* pelvic.	235
Test 23. MRI** pelvic.	236
Test 24. MRI ovarian.	236
Test 25. MRI DIE.	236
Test 26. MRI posterior DIE.	236
Test 27. MRI* posterior DIE.	236
Test 28. MRI USL.	236
Test 29. MRI* USL.	236
Test 30. MRI RVS.	236
Test 31. MRI vaginal.	236
Test 32. MRI* vaginal.	236
Test 33. MRI POD.	236
Test 34. MRI* POD.	236
Test 35. MRI anterior DIE.	236
Test 36. MRI rectosigmoid.	237
Test 37. MRI* rectosigmoid.	237
Test 38. MDCT-e rectosigmoid.	237
Test 39. MDCT-e bowel [ileum - rectum].	237
Test 40. 18FDG PET-CT pelvic.	237
Test 41. DCBE DIE.	237
Test 42. DCBE rectosigmoid.	237
Test 43. MRI pelvic1.	237
ADDITIONAL TABLES	237
APPENDICES	248
CONTRIBUTIONS OF AUTHORS	254
DECLARATIONS OF INTEREST	254
SOURCES OF SUPPORT	254
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	254
NOTES	255
INDEX TERMS	255

[Diagnostic Test Accuracy Review]

Imaging modalities for the non-invasive diagnosis of endometriosis

Vicki Nisenblat¹, Patrick MM Bossuyt², Cindy Farquhar³, Neil Johnson¹, M Louise Hull¹

¹Discipline of Obstetrics and Gynaecology, School of Medicine, Robinson Research Institute, The University of Adelaide, Adelaide, Australia. ²Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands. ³Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand

Contact address: Vicki Nisenblat, Discipline of Obstetrics and Gynaecology, School of Medicine, Robinson Research Institute, The University of Adelaide, Level 6, Medical School North,, Frome Rd, Adelaide, SA, 5005, Australia. vnisenblat@gmail.com.

Editorial group: Cochrane Gynaecology and Fertility Group.

Publication status and date: New, published in Issue 2, 2016.

Citation: Nisenblat V, Bossuyt PMM, Farquhar C, Johnson N, Hull ML. Imaging modalities for the non-invasive diagnosis of endometriosis. *Cochrane Database of Systematic Reviews* 2016, Issue 2. Art. No.: CD009591. DOI: [10.1002/14651858.CD009591.pub2](https://doi.org/10.1002/14651858.CD009591.pub2).

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

About 10% of women of reproductive age suffer from endometriosis. Endometriosis is a costly chronic disease that causes pelvic pain and subfertility. Laparoscopy, the gold standard diagnostic test for endometriosis, is expensive and carries surgical risks. Currently, no non-invasive tests that can be used to accurately diagnose endometriosis are available in clinical practice. This is the first review of diagnostic test accuracy of imaging tests for endometriosis that uses Cochrane methods to provide an update on the rapidly expanding literature in this field.

Objectives

- To provide estimates of the diagnostic accuracy of imaging modalities for the diagnosis of pelvic endometriosis, ovarian endometriosis and deeply infiltrating endometriosis (DIE) versus surgical diagnosis as a reference standard.
- To describe performance of imaging tests for mapping of deep endometriotic lesions in the pelvis at specific anatomical sites.

Imaging tests were evaluated as replacement tests for diagnostic surgery and as triage tests that would assist decision making regarding diagnostic surgery for endometriosis.

Search methods

We searched the following databases to 20 April 2015: MEDLINE, CENTRAL, EMBASE, CINAHL, PsycINFO, Web of Science, LILACS, OAlster, TRIP, ClinicalTrials.gov, MEDION, DARE, and PubMed. Searches were not restricted to a particular study design or language nor to specific publication dates. The search strategy incorporated words in the title, abstracts, text words across the record and medical subject headings (MeSH).

Selection criteria

We considered published peer-reviewed cross-sectional studies and randomised controlled trials of any size that included prospectively recruited women of reproductive age suspected of having one or more of the following target conditions: endometrioma, pelvic endometriosis, DIE or endometriotic lesions at specific intrapelvic anatomical locations. We included studies that compared the diagnostic test accuracy of one or more imaging modalities versus findings of surgical visualisation of endometriotic lesions.

Data collection and analysis

Two review authors independently collected and performed a quality assessment of data from each study. For each imaging test, data were classified as positive or negative for surgical detection of endometriosis, and sensitivity and specificity estimates were calculated. If two or

more tests were evaluated in the same cohort, each was considered as a separate data set. We used the bivariate model to obtain pooled estimates of sensitivity and specificity when sufficient data sets were available. Predetermined criteria for a clinically useful imaging test to replace diagnostic surgery included sensitivity $\geq 94\%$ and specificity $\geq 79\%$. Criteria for triage tests were set at sensitivity $\geq 95\%$ and specificity $\geq 50\%$, ruling out the diagnosis with a negative result (SnNout test - if sensitivity is high, a negative test rules out pathology) or at sensitivity $\geq 50\%$ with specificity $\geq 95\%$, ruling in the diagnosis with a positive result (SpPin test - if specificity is high, a positive test rules in pathology).

Main results

We included 49 studies involving 4807 women: 13 studies evaluated pelvic endometriosis, 10 endometriomas and 15 DIE, and 33 studies addressed endometriosis at specific anatomical sites. Most studies were of poor methodological quality. The most studied modalities were transvaginal ultrasound (TVUS) and magnetic resonance imaging (MRI), with outcome measures commonly demonstrating diversity in diagnostic estimates; however, sources of heterogeneity could not be reliably determined. No imaging test met the criteria for a replacement or triage test for detecting pelvic endometriosis, albeit TVUS approached the criteria for a SpPin triage test. For endometrioma, TVUS (eight studies, 765 participants; sensitivity 0.93 (95% confidence interval (CI) 0.87, 0.99), specificity 0.96 (95% CI 0.92, 0.99)) qualified as a SpPin triage test and approached the criteria for a replacement and SnNout triage test, whereas MRI (three studies, 179 participants; sensitivity 0.95 (95% CI 0.90, 1.00), specificity 0.91 (95% CI 0.86, 0.97)) met the criteria for a replacement and SnNout triage test and approached the criteria for a SpPin test. For DIE, TVUS (nine studies, 12 data sets, 934 participants; sensitivity 0.79 (95% CI 0.69, 0.89) and specificity 0.94 (95% CI 0.88, 1.00)) approached the criteria for a SpPin triage test, and MRI (six studies, seven data sets, 266 participants; sensitivity 0.94 (95% CI 0.90, 0.97), specificity 0.77 (95% CI 0.44, 1.00)) approached the criteria for a replacement and SnNout triage test. Other imaging tests assessed in small individual studies could not be statistically evaluated.

TVUS met the criteria for a SpPin triage test in mapping DIE to uterosacral ligaments, rectovaginal septum, vaginal wall, pouch of Douglas (POD) and rectosigmoid. MRI met the criteria for a SpPin triage test for POD and vaginal and rectosigmoid endometriosis. Transrectal ultrasonography (TRUS) might qualify as a SpPin triage test for rectosigmoid involvement but could not be adequately assessed for other anatomical sites because heterogeneous data were scant. Multi-detector computerised tomography enema (MDCT-e) displayed the highest diagnostic performance for rectosigmoid and other bowel endometriosis and met the criteria for both SpPin and SnNout triage tests, but studies were too few to provide meaningful results.

Diagnostic accuracies were higher for TVUS with bowel preparation (TVUS-BP) and rectal water contrast (RWC-TVS) and for 3.0TMRI than for conventional methods, although the paucity of studies precluded statistical evaluation.

Authors' conclusions

None of the evaluated imaging modalities were able to detect overall pelvic endometriosis with enough accuracy that they would be suggested to replace surgery. Specifically for endometrioma, TVUS qualified as a SpPin triage test. MRI displayed sufficient accuracy to suggest utility as a replacement test, but the data were too scant to permit meaningful conclusions. TVUS could be used clinically to identify additional anatomical sites of DIE compared with MRI, thus facilitating preoperative planning. Rectosigmoid endometriosis was the only site that could be accurately mapped by using TVUS, TRUS, MRI or MDCT-e. Studies evaluating recent advances in imaging modalities such as TVUS-BP, RWC-TVS, 3.0TMRI and MDCT-e were observed to have high diagnostic accuracies but were too few to allow prudent evaluation of their diagnostic role. In view of the low quality of most of the included studies, the findings of this review should be interpreted with caution. Future well-designed diagnostic studies undertaken to compare imaging tests for diagnostic test accuracy and costs are recommended.

PLAIN LANGUAGE SUMMARY

Imaging tests for the non-invasive diagnosis of endometriosis

Review question

How accurate are imaging tests in detecting endometriosis? Can any imaging test be accurate enough to replace or reduce the need for surgery in the diagnosis of endometriosis?

Background

Women with endometriosis have endometrial tissue (the tissue that lines the womb and is shed during menstruation) growing outside the womb within the pelvis, causing chronic abdominal pain and difficulty conceiving. Currently, the only reliable way of diagnosing endometriosis is to perform laparoscopic surgery and visualise the endometrial deposits inside the abdomen. Because surgery is risky and expensive, imaging tests have been assessed for their ability to detect endometriosis non-invasively. An accurate imaging test could lead to the diagnosis of endometriosis without the need for surgery, or it could reduce the need for surgery, so only women who were most likely to have endometriosis would require it. Furthermore, if imaging tests could accurately predict the location of endometriotic lesions, surgeons would have the information they need to plan and improve their surgical approach. Other non-invasive ways of diagnosing endometriosis by using urine, blood and endometrial and combination tests have been evaluated in separate Cochrane reviews from this series.

Study characteristics

Imaging modalities for the non-invasive diagnosis of endometriosis (Review)

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Evidence included in this review is current to April 2015. We included 49 studies involving 4807 participants. Thirteen studies evaluated pelvic endometriosis, 10 studies ovarian endometrioma, 15 studies deep endometriosis (endometriosis deeply situated in tissues in the pelvis) and 33 studies endometriosis at specific sites within the pelvic cavity. All studies included women of reproductive age who were undergoing diagnostic surgery because they had symptoms of endometriosis.

Key results

None of the imaging methods was accurate enough to provide this information on overall pelvic endometriosis. Transvaginal ultrasound identified ovarian endometriosis with enough accuracy to help surgeons determine whether surgery was needed, and magnetic resonance imaging (MRI) was sufficiently accurate to replace surgery in diagnosing endometrioma but was evaluated in only a small number of studies. Other imaging tests were assessed in small individual studies and could not be evaluated in a meaningful way. Transvaginal ultrasound could be used to locate more anatomical sites of deep endometriosis when compared with MRI, helping surgeons better plan an operative procedure. Endometriosis in the lower bowel appears to be relatively accurately identified by both transvaginal and transrectal ultrasound, by MRI and by multi-detector computerised tomography enema. New types of ultrasound and MRI show a lot of promise in detecting endometriosis but studies are too few to clearly show their diagnostic value.

Quality of the evidence

Generally the studies were of low methodological quality, and most imaging techniques were assessed by only a small number of studies. Differences between studies involved how they were run, groups of women studied, ways imaging tests were performed and how surgery was undertaken.

Future research

Additional high-quality research is needed to accurately evaluate the diagnostic potential of non-invasive imaging tests for endometriosis.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings table: diagnostic tests for endometriosis

Review question	What is the diagnostic accuracy of the imaging tests in detecting endometriosis?		Pelvic endometriosis (any site and depth of invasion)
			Ovarian endometriosis
			DIE
Importance	A simple and reliable non-invasive test for endometriosis with the potential to replace laparoscopy or to triage patients to reduce surgery would minimise surgical risk and reduce diagnostic delay		
Participants	Women of reproductive age (1) with suspected endometriosis and/or (2) with persistent ovarian mass and/or (3) undergoing infertility workup		
Settings	Hospitals (public or private of any level); outpatient clinics (general gynaecology, reproductive medicine, pelvic pain) and/or radiology departments		
Reference standard	Visualisation of endometriosis at surgery (laparoscopy or laparotomy) with or without histological confirmation		
Study design	Cross-sectional of 'single-gate' design (n = 28) or 'two-gate' design (n = 1); prospective enrolment; 1 study could assess more than 1 test and/or more than 1 type of endometriosis		
Risk of bias and applicability concerns	Overall judgement	Poor quality of most studies (only 1 study had 'low risk' assessment in all 4 domains; Thomeer 2014)	
	Patient selection bias	High risk: 13 studies; unclear risk: 6 studies; low risk: 10 studies	
	Index test interpretation bias	High risk: 7 studies; unclear risk: 7 studies; low risk: 15 studies	
	Reference standard interpretation bias	High risk: 6 studies; unclear risk: 16 studies; low risk: 7 studies	
	Flow and timing se-	High risk: 9 studies; unclear risk: 2 studies; low risk: 18 studies	

lection
bias

Applica-
bility con-
cerns

Concerns regarding patient selection: high concern - 1 study, unclear concern - 0 studies, low concern - 28 studies

Concerns regarding index test: high concern - 0 studies, unclear concern - 0 studies, low concern - 29 studies

Concerns regarding reference standard: high concern - 0 studies, unclear concern - 0 studies, low concern - 29 studies

**Diagnos-
tic thresh-
olds**

Replacement test: sensitivity $\geq 94\%$; specificity $\geq 79\%$

SnNout triage test: sensitivity $\geq 95\%$; specificity $\geq 50\%$

SpPin triage test: sensitivity $\geq 50\%$; specificity $\geq 95\%$

Approaching criteria for 1 of the above tests: diagnostic estimates within 5% of set thresholds

Target condition	Test	N of participants; N of studies; N of data sets	Pooled estimates (95% CI)	Outcomes				Implications
				True posi- tives (en- dometrio- sis)	False pos- itives (in- correctly classi- fied as en- dometrio- sis)	False neg- atives (in- correctly classified as dis- ease-free)	True nega- tives (dis- ease-free)	
<u>Pelvic en- dometrio- sis (13 studies, 1535 par- ticipants)</u>	TVUS	1222 participants in 5 studies	Sens = 0.65 (0.27 to 1.00) Spec = 0.95 (0.89 to 1.00) <u>Meta-analysis of 4 stud- ies after removing 1 out- lier study</u> <u>Sens = 0.79 (0.36 to 1.00)</u> <u>Spec = 0.91 (0.74 to 1.00)</u>	257	24	372	569	Approaches the criteria for a SpPin triage test when 1 outlier study was excluded. Wide confidence intervals (CIs)
	MRI	303 participants in 7 studies; 396 participants in 10 data sets	Sens = 0.79 (0.70 to 0.88) Spec = 0.72 (0.51 to 0.92)	253	21	70	52	Neither replacement nor triage test criteria met Observation: 3.0T MRI (2 studies) demonstrated highest diagnostic accuracy

	¹⁸ FGD PET-CT	10 participants in 1 study	Not available ^a	0	0	9	1	Insufficient evidence to allow meaningful conclusions
<u>Ovari-an en-dometrio-sis (10 studies, 852 participants)</u>	TVUS	765 participants in 8 studies	Sens = 0.93 (0.87 to 0.99) Spec = 0.96 (0.92 to 0.99)	182	28	16	539	Meets the criteria for a SpPin triage test and approaches the criteria for a replacement and SnNout triage test Observation: Studies published after 2006 (4 out of 5 studies) demonstrated highest diagnostic accuracy
	TRUS	92 participants in 1 study	Not available ^b	32	13	4	43	Insufficient evidence to allow meaningful conclusions
	MRI	179 participants in 3 studies	Sens = 0.95 (0.90 to 1.00) Spec = 0.91 (0.86 to 0.97)	72	9	4	94	Meets the criteria for a replacement and SnNout triage test, approaches the criteria for a SpPin triage test Observation: 3.0T MRI (2 studies) demonstrated highest diagnostic accuracy Insufficient evidence to allow meaningful conclusions
<u>DIE/Posterior DIE (15 studies, 1493 participants)</u>	TVUS	934 participants in 9 studies; 1383 participants in 12 data sets	Sens = 0.79 (0.69 to 0.89) Spec = 0.94 (0.88 to 1.00)	435	51	128	769	Approaches the criteria for a SpPin triage test Observation: TVUS-BP (1 study) demonstrated highest diagnostic accuracy
	MRI	266 participants in 6 studies; 289 participants in 7 data sets	Sens = 0.94 (0.90 to 0.97) Spec = 0.77 (0.44 to 1.00)	210	11	9	59	Approaches the criteria for a replacement and SnNout triage test Observation: 3.0T MRI (2 studies) and MRI jelly method (1 study) demonstrated highest diagnostic accuracy
	DCBE	69 participants in 1 study	Not available ^c	24	0	43	2	Insufficient evidence to allow meaningful conclusions

^aFor FGD PET-CT in pelvic endometriosis, diagnostic estimates were sensitivity = 0.00 (0.00 to 0.34); specificity = 1.00 (0.03 to 1.00)

^bFor TRUS in ovarian endometriosis, diagnostic estimates were sensitivity = 0.89 (0.74 to 0.97); specificity = 0.77 (0.64 to 0.87)

^cFor DCBE in DIE, diagnostic estimates were sensitivity = 0.36 (0.24 to 0.48); specificity = 1.00 (0.16 to 1.00)

Summary of findings 2. Summary of findings table: surgical mapping of endometriosis to specific anatomical sites

Review question	What is the diagnostic performance of the imaging tests in mapping deep endometriotic lesions in the pelvis at specific anatomical sites?		USL endometriosis
			RVS endometriosis
			Vaginal wall endometriosis
			POD obliteration
			Anterior DIE
			RS/Bowel endometriosis
Importance	Ability to diagnose DIE at specific anatomical sites at preoperative assessment helps optimise planning of surgery or guides referral to the most appropriate practice, with the potential to improve treatment outcomes		
Participants	Women of reproductive age with suspected endometriosis or specifically suspected DIE		
Settings	Hospitals (public or private of any level); outpatient clinics (general gynaecology, reproductive medicine, pelvic pain) and/or radiology departments		
Reference standard	Visualisation of endometriosis at surgery (laparoscopy or laparotomy) with or without histological confirmation		
Study design	Cross-sectional of 'single-gate' design (n = 33); prospective enrolment; 1 study could assess more than 1 test and/or more than 1 site of endometriosis		
Risk of bias and applicability concerns	Overall judgement	Poor quality of most studies (only 1 study had 'low risk' assessment in all 4 domains; Thomeer 2014)	
	Patient selection bias	High risk: 16 studies; unclear risk: 6 studies; low risk: 11 studies	
	Index test interpretation bias	High risk: 8 studies; unclear risk: 4 studies; low risk: 21 studies	
	Reference standard interpretation bias	High risk: 14 studies; unclear risk: 14 studies; low risk: 5 studies	
	Flow and timing selection bias	High risk: 8 studies; unclear risk: 3 studies; low risk: 22 studies	

Applicability concerns		Concerns regarding patient selection: high concern - 0 studies, unclear concern - 0 studies, low concern - 33 studies						
		Concerns regarding index test: high concern - 0 studies, unclear concern - 0 studies, low concern - 33 studies						
		Concerns regarding reference standard: high concern - 0 studies, unclear concern - 0 studies, low concern - 33 studies						
Diagnostic thresholds	Replacement test: sensitivity ≥ 94%; specificity ≥ 79%							
	SnNout triage test: sensitivity ≥ 95%; specificity ≥ 50%							
	SpPin triage test: sensitivity ≥ 50%; specificity ≥ 95%							
	Approaching criteria for 1 of the above tests: diagnostic estimates within 5% of set thresholds							
Target condition	Test	N of participants; N of studies; N of data sets	Pooled estimates (95% CI)	Outcomes				Implications
				True positives (endometriosis)	False positives (incorrectly classified as endometriosis)	False negatives (incorrectly classified as disease-free)	True negatives (disease-free)	
<u>USL endometriosis (11 studies, 997 participants)</u>	TVUS	751 participants in 7 studies	Sens = 0.64 (0.50 to 0.79) Spec = 0.97 (0.93 to 1.00)	136	18	63	534	Meets the criteria for a SpPin triage test Observation: TVUS-BP (1 study) demonstrated the highest diagnostic accuracy
	TRUS	232 participants in 2 studies	Sens = 0.52 (0.29 to 0.74) Spec = 0.94 (0.86 to 1.00)	48	8	45	131	Approach the criteria for a SpPin triage test Wide CIs Insufficient evidence to allow meaningful conclusions
	MRI	199 participants in 4 studies 221 participants in 5 data sets	Sens = 0.86 (0.80 to 0.92) Spec = 0.84 (0.68 to 1.00)	136	13	22	50	Criteria for a triage test not met Wide CIs Observation: 3.0T MRI (1 out of 2 studies) demonstrated the highest diagnostic accuracy

<u>RVS endometrios (12 studies, 1215 participants)</u>	TVUS	983 participants in 10 studies 1073 participants in 11 data sets	Sens = 0.88 (0.82 to 0.94) Spec = 1.00 (0.98 to 1.00)	263	10	59	741	Meets the criteria for a SpPin triage test Observation: TVUS-BP (3 studies) and RWC-TVS (1 study) demonstrated the highest diagnostic accuracy
	TRUS	232 participants in 2 studies	Sens = 0.78 (0.51 to 1.00) Spec = 0.96 (0.89 to 1.00)	35	8	10	179	Meets the criteria for a SpPin triage test Insufficient evidence to allow meaningful conclusions
	MRI	288 participants in 3 studies	Sens = 0.81 (0.70 to 0.93) Spec = 0.86 (0.78 to 0.95)	96	23	22	147	Criteria for a triage test not met Insufficient evidence to allow meaningful conclusions
<u>Vaginal wall endometrios (10 studies, 981 participants)</u>	TVUS	679 participants in 6 studies	Sens = 0.57 (0.21 to 0.94) Spec = 0.99 (0.96 to 1.00)	70	11	44	554	Meets the criteria for a SpPin triage test Wide CIs Observation: tg-TVUS (1 study) demonstrated the highest diagnostic accuracy
	TRUS	232 participants in 2 studies	Sens = 0.39 (0.08 to 0.70) Spec = 1.00 (1.00 to 1.00)	18	0	28	186	Criteria for a triage test not met Wide CIs Insufficient evidence to allow meaningful conclusions
	MRI	248 participants in 4 studies 271 participants in 5 data sets	Sens = 0.77 (0.67 to 0.88) Spec = 0.97 (0.92 to 1.00)	48	11	14	198	Meets the criteria for a SpPin triage test Observation: 3.0T MRI (1 study) and 3D-MRI demonstrated the highest diagnostic accuracy
<u>POD obliteration (11 studies, 909</u>	TVUS	755 participants in 6 studies	Sens = 0.83 (0.77 to 0.88) Spec = 0.97 (0.95 to 0.99)	152	17	32	554	Meets the criteria for a SpPin triage test

<u>partici-</u> <u>pants)</u>								Observation: TVUS-BP (2 stud- ies) demonstrated the highest diagnostic accuracy
	MRI	154 participants in 5 studies 177 participants in 6 data sets	Sens = 0.90 (0.76 to 1.00) Spec = 0.98 (0.89 to 1.00)	84	3	12	78	Meets the criteria for a SpPin triage test and approaches the criteria for a SnNout triage test Observation: 3.0T MRI (3 stud- ies) demonstrated the highest diagnostic accuracy
<u>Anterior</u> <u>DIE</u> (3 studies, 330 partic- ipants)	TVUS	289 participants in 2 studies	Sens = 0.41 (0.00 to 0.81) Spec = 1.00 (1.00 to 1.00)	11	0	16	262	Criteria for a triage test not met Wide CIs Insufficient evidence to allow meaningful conclusions
	MRI	41 participants in 1 study	Not available ^a	6	0	2	33	Insufficient evidence to allow meaningful conclusions
<u>Rectosig-</u> <u>moid en-</u> <u>dometrio-</u> <u>sis</u> (21 stud- ies, 2222 partici- pants)	TVUS	1616 participants in 14 studies 1817 participants in 15 data sets	Sens = 0.90 (0.82 to 0.97) Spec = 0.96 (0.94 to 0.99)	648	47	100	1022	Meets the criteria for a SpPin triage test and approaches the criteria for a SnNout triage test Observation: TVUS-BP (2 stud- ies) and RWC-TVS (2 studies) demonstrated the highest diag- nostic accuracy
	TRUS	330 participants in 4 studies	Sens = 0.91 (0.85 to 0.98) Spec = 0.96 (0.91 to 1.00)	137	8	13	172	Meets the criteria for a SpPin triage test and approaches the criteria for a SnNout triage test
	MRI	612 participants in 6 studies 635 participants in 7 data sets	Sens = 0.92 (0.86 to 0.99) Spec = 0.96 (0.93 to 0.98)	352	11	30	242	Meets the criteria for a SpPin triage test and approaches the criteria for a SnNout triage test Observation: MRI jelly method (1 study) and 3.0T MRI (1 study) demonstrated the highest diag- nostic accuracy
	MDCT-e	389 participants in 3 studies	Sens = 0.98 (0.94 to 1.00) Spec = 0.99 (0.97 to 1.00)	241	1	6	141	Meets the criteria for a SpPin test and a SnNout triage test

								Insufficient evidence to allow meaningful conclusions
	DCBE	106 participants in 2 studies	Sens = 0.56 (0.32 to 0.80) Spec = 0.77 (0.41 to 1.00)	45	6	35	20	Criteria for a triage test not met Wide CIs Insufficient evidence to allow meaningful conclusions
<u>Bowel (ileum - rectum) endometriosis (4 studies, 412 participants)</u>	TVUS	314 participants in 3 studies	Sens = 0.89 (0.81 to 0.97) Spec = 0.96 (0.91 to 1.00)	135	7	16	156	Meets the criteria for a SpPin triage test Observation: TVUS, non-modified method (1 study) demonstrated highest diagnostic estimates Insufficient evidence to allow meaningful conclusions
	TRUS	134 participants in 1 study	Not available ^b	72	0	3	59	Insufficient evidence to allow meaningful conclusions
	MDCT-e	194 participants in 2 studies	Sens = 0.98 (0.92 to 1.00) Spec = 1.00 (1.00 to 1.00)	124	0	3	67	Meets the criteria for a SpPin test and a SnNout triage test Insufficient evidence to allow meaningful conclusions

^aFor MRI in anterior DIE, diagnostic estimates were sensitivity = 0.75 (0.35 to 0.97); specificity = 1.00 (0.89 to 1.00)

^bFor TRUS in bowel endometriosis, diagnostic estimates were sensitivity = 0.96 (0.89 to 0.99); specificity = 1.00 (0.94 to 1.00)

BACKGROUND

Target condition being diagnosed

Endometriosis

Endometriosis is defined as an inflammatory condition characterised by endometrium-like tissue at sites outside the uterus ([Johnson and Hummelshoj 2013](#)). Endometriotic lesions can be found at different locations, including the pelvic peritoneum and the ovary, or can penetrate pelvic structures below the surface of the peritoneum as deeply infiltrating endometriosis (DIE). Each of these types of endometriosis is thought to represent a separate clinical entity, but different types can co-exist in the same woman. Pelvic endometriosis is defined as the presence of any endometrial tissue within the pelvic cavity, including the peritoneum, within any of the pelvic organs and inside the pouch of Douglas (POD). Ovarian endometriosis, an endometrioma, is defined as an ovarian cyst lined by endometrial tissue; it appears as ovarian masses of varying size. Endometriomas are identified more easily by imaging or by pelvic examination than are other forms of endometriosis; however, discrimination of benign ovarian endometriosis from other types of ovarian tumours can be challenging. DIE is defined as endometriotic tissue that penetrates the retroperitoneal space for a distance of 5 mm or more ([Koninckx 1991](#)) and may be present in multiple locations, involving anterior or posterior pelvic compartments, or both. Posterior DIE, a multi-focal disease that may affect a variety of anatomical sites, represents the most common type of DIE ([Kinkel 2006](#)). The most typical sites of DIE include uterosacral ligaments (USL), rectovaginal septum (RVS), vaginal wall, POD and bowel, predominantly below the rectosigmoid junction. Anterior DIE corresponds to disease involving the anterior pouch or bladder and is much less common. Rarely, endometriotic implants can be found at more distant sites, including lung, liver, pancreas and operative scars, with consequent variation in presenting symptoms.

Endometriosis afflicts 10% of women of reproductive age, causing dysmenorrhoea (painful periods), dyspareunia (painful intercourse), chronic pelvic pain and infertility ([Vigano 2004](#)). The clinical presentation can vary from asymptomatic and unexplained infertility to severe dysmenorrhoea and chronic pain. Symptoms can occur with bowel or urinary symptoms, an abnormal pelvic examination or the presence of a pelvic mass; however, no symptom is specific to endometriosis. Prevalence of endometriosis in the symptomatic population is reported as 35% to 50% ([Giudice 2004](#)).

Women with endometriosis are at increased risk of developing several cancers ([Somigliana 2006](#)) and autoimmune disorders ([Sinaii 2002](#)). The presence of disease is associated with changes in immune response, vascularisation, neural function, peritoneal environment and eutopic endometrium, suggesting that endometriosis is a systemic, rather than a localised, condition ([Giudice 2004](#)). Endometriosis has a profound effect on psychological and social well-being and imposes a substantial economic burden on society. Women with endometriosis incur significant direct medical costs from diagnostic and therapeutic surgeries, hospital admissions and fertility treatments; however, these costs are superseded by indirect costs of endometriosis, including absenteeism and loss of productivity ([Gao 2006](#); [Simoens 2012](#)). In the United States, the financial burden of endometriosis is estimated at US \$12,419 per woman ([Simoens 2012](#)).

Although the pathogenesis of endometriosis has not been fully elucidated, it is commonly thought that endometriosis occurs when endometrial tissue contained within menstrual fluid flows retrogradely through the fallopian tubes and implants at an ectopic site within the pelvic cavity ([Sampson 1927](#)). However, this theory does not explain the fact that although retrograde menstruation is seen in up to 90% of women, only 10% of women develop endometriosis ([Halme 1984](#)). Evidence suggests that a variety of environmental, immunological and hormonal factors are associated with endometriosis ([Vigano 2004](#)), and genetic loci that confer risk of endometriosis have been identified ([Nyholt 2012](#)). The relative contributions of these and other causal factors remain to be elucidated.

Although it is impossible to time the onset of disease, on average, women have a six- to 12-year history of symptoms before obtaining a surgical diagnosis of endometriosis, which indicates considerable diagnostic delay ([Matsuzaki 2006](#)). Untreated endometriosis is associated with reduced quality of life and contributes to outcomes such as depression, inability to work, sexual dysfunction and missed opportunities for motherhood ([Gao 2006](#)).

Treatment of endometriosis

No cure for endometriosis is known. Treatment options include expectant management, pharmacological (hormonal) therapy and surgery ([Johnson and Hummelshoj 2013](#)). Treatment is individualised, taking into consideration the therapeutic goal (pain relief or subfertility) and the location of the disease. Current pharmacological therapies such as the combined oral contraceptive pill, progestogens, weak androgens and gonadotropin-releasing hormone (GnRH) agonists and antagonists act to reduce the effects of oestrogen on endometrial tissues and to suppress menstruation. These drugs can ameliorate symptoms of dysmenorrhoea and chronic pelvic pain, but they are associated with side effects such as breast discomfort, irritability, androgenic symptoms and bone loss. Surgical excision of endometriotic lesions can reduce pain and improve fertility, but is associated with high recurrence rates of 40% to 50% at five years post surgery ([Guo 2009](#); [Duffy 2014](#)). Early treatment of individuals with endometriosis improves pain levels and physical and psychological functioning. Furthermore, improvements in management of menstruation (use of the Mirena coil and continuous use of the combined contraceptive pill) and fertility preservation (oocyte vitrification) raise the possibility of suppressing the progression of endometriosis and prospectively managing subfertility among endometriosis sufferers. The potential success of these preventative strategies is dependent on an accurate and early diagnosis. A major impediment to earlier and more efficacious treatment of this disease is diagnostic delay due to the invasive nature of standard diagnostic tests ([Dmowski 1997](#)).

Diagnosis of endometriosis

Clinical history and pelvic examination can raise the possibility of a diagnosis of endometriosis, but heterogeneity in clinical presentation, high prevalence of asymptomatic endometriosis (2% to 50%) and poor association between presenting symptoms and severity of the disease contribute to the difficulty involved in obtaining a reliable diagnosis of endometriosis based solely on presenting symptoms ([Spaczynski 2003](#); [Fauconnier 2005](#); [Ballard 2008](#)). Although an abnormal pelvic examination correlates

with the presence of endometriosis on laparoscopy in 70% to 90% of cases (Ling 1999), the differential diagnosis for most positive physical findings is wide. Furthermore, a normal clinical examination does not exclude endometriosis, as laparoscopically proven disease has been diagnosed in more than 50% of women with a clinically normal pelvic examination (Eskenazi 2001). A variety of tests utilising pelvic imaging, blood markers, eutopic endometrium characteristics, urinary markers or peritoneal fluid components have been suggested as diagnostic measures for endometriosis. Although large numbers of the reported markers have distinguished women with and without endometriosis in small pilot studies, many have not shown convincing potential as a diagnostic test when evaluated in larger studies by different research groups. The diagnostic value of these tests has not been fully systematically evaluated and summarised by Cochrane methods. Currently, no simple non-invasive test for the diagnosis of endometriosis is routinely implemented in clinical practice.

Surgical diagnostic procedures for endometriosis include laparoscopy (minimal access surgery) or laparotomy (open surgery via an abdominal incision). Over the past several decades, laparoscopy has become an increasingly common procedure that has largely replaced traditional open surgery among women suspected of having endometriosis (Yeung 2009). Laparoscopy confers significant advantages over laparotomy, creating fewer complications and shorter recovery times. Furthermore, a magnified view at laparoscopy allows better visualisation of the peritoneal cavity. Despite continuing controversy in the literature with regard to the superiority of one surgical modality over another for treating women with pelvic disease, laparoscopy is the preferred technique for evaluating the pelvis and abdomen and for treating individuals with benign conditions such as ovarian endometrioma (Medeiros 2009). Surgery is also the only currently accepted way to determine the extent and severity of endometriosis. Several classification systems have been suggested for endometriosis (Batt 2003; Chapron 2003a; Martin 2006; Adamson 2008), but most researchers and clinicians use the revised American Society for Reproductive Medicine (rASRM) classification, which is internationally accepted as a respected currently available tool for objective assessment of the disease (American Society for Reproductive Medicine 1997). The rASRM classification system considers appearance, size and depth of peritoneal or ovarian implants and adhesions visualised during laparoscopy (Table 1) and allows uniform documentation of the extent of disease. Unfortunately, this classification system has little value in clinical practice because of lack of correlation between laparoscopic staging, severity of symptoms and response to treatment (Vercellini 1996; Guzick 1997; Chapron 2003b). A recent endeavour to attain consensus around the optimal classification for endometriosis has been undertaken by the World Endometriosis Society (Johnson 2015).

The European Society for Human Reproduction and Embryology (ESHRE) Special Interest Group for Endometriosis stated in its guidelines for the diagnosis and treatment of endometriosis that for women presenting with symptoms suggestive of endometriosis, a definitive diagnosis of most forms of endometriosis requires visual inspection of the pelvis at laparoscopy as the 'gold standard' investigation (Kennedy 2005). Currently, the visual or histological identification of endometriotic tissue in the pelvic cavity during surgery is not just the best available but the only diagnostic test for endometriosis that is used routinely in clinical practice.

Disadvantages of laparoscopic surgery include and are not limited to high cost, need for general anaesthesia and potential for adhesion formation post procedure. Laparoscopy has been associated with 2% risk of injury to pelvic organs, 0.001% risk of damage to a major blood vessel and a mortality rate of 0.0001% (Chapron 2003c). Only one-third of women who undergo a laparoscopic procedure will receive a diagnosis of endometriosis; therefore, many disease-free women are unnecessarily exposed to surgical risk (Frishman 2006).

The validity of laparoscopy as a reference test for endometriosis has been assessed as highly dependent on the skills of the surgeon. The diagnostic accuracy of laparoscopic visualisation has been compared with histological confirmation in a sole systematic review; 94% sensitivity and 79% specificity have been reported (Wykes 2004). Subsequent studies suggested that incorporation of histological verification into the diagnosis of endometriosis may improve diagnostic accuracy (Marchino 2005; Almeida Filho 2008; Stegmann 2008), but these papers have not been systematically reviewed. The clinical significance of histological verification remains debatable, and a diagnosis based on visual findings can be considered reliable with accurate inspection of the abdominal cavity by properly trained experienced surgeons (Redwine 2003). Furthermore, excised potentially endometriotic tissues are rarely serially sectioned in clinical practice, and small lesions can be missed by pathologists in cases of mild disease. Thus sampling inconsistencies are likely to influence the accuracy of histological reporting.

Summary

A diagnostic test in place of surgery would reduce associated surgical risks, increase diagnostic accessibility and improve treatment outcomes. The need for an accurate and non-invasive diagnostic test for endometriosis continues to encourage extensive research in the field and was endorsed at the international consensus workshop at the 10th World Congress of Endometriosis in 2008 (Rogers 2009). Although multiple markers and imaging techniques have been explored as diagnostic tests for endometriosis, none of them have been implemented routinely in clinical practice, and many have not been subject to systematic review.

Index test(s)

This review assesses the diagnostic imaging techniques that have been proposed as non-invasive tests for the diagnosis of endometriosis (Table 2) as part of the review series on non-invasive diagnostic tests for endometriosis. The other reviews from this series include 'Blood biomarkers for the non-invasive diagnosis of endometriosis', 'Endometrial biomarkers for the non-invasive diagnosis of endometriosis', 'Urinary biomarkers for the non-invasive diagnosis of endometriosis' and 'Combination of the non-invasive tests for the diagnosis of endometriosis', which is the summary review for this series.

The definition of 'non-invasive' varies between medical dictionaries, but the term refers to a procedure that does not involve penetration of skin or physical entrance into the body (McGraw-Hill Dictionary of Medicine 2006; The Gale Encyclopedia of Medicine 2008). Although some imaging tests are associated with an intracavitary approach (e.g. transvaginal, transrectal) and therefore are invasive by this definition, when compared with

diagnostic surgery for endometriosis, these tests are generally considered to be 'non-invasive' or 'minimally invasive'. For the purpose of these reviews, we will define all tests that do not involve anaesthesia and surgery as 'non-invasive'.

Advantages of using imaging tests for the diagnosis of endometriosis include that they are minimally invasive, readily available and more acceptable to women; provide a rapid result; and are more cost-effective when compared with surgery. However, imaging testing is dependent on the skills of the operator and the ability of women to access appropriate radiology services. At this point in time, all imaging modalities have been assessed in a limited number of small studies, which vary in the type of imaging methods used and the anatomical locations evaluated.

Magnetic resonance imaging (MRI) and ultrasonography (US) (which includes transabdominal, transvaginal and transrectal approaches) are the most widely reported diagnostic modalities for endometriosis. A systematic review that primarily summarised the diagnostic performance of ultrasound for endometriosis-associated ovarian masses (endometriomas) concluded that transvaginal ultrasound (TVUS) has clinical utility in differentiating endometriomas from other types of ovarian cysts (Moore 2002). This review concentrated on studies that used transabdominal and transvaginal US with or without Doppler and did not include reports on other forms of ultrasound, nor did it evaluate non-ovarian forms of endometriosis. Studies that evaluated the ability of ultrasound to detect endometriotic implants at other pelvic sites reported varying degrees of accuracy for deep endometriotic lesions and failure to detect small lesions and pelvic adhesions (Kinkel 2006). Because of high costs and limited availability, MRI is not frequently implemented in routine clinical practice; however, a growing number of studies suggest that it has a role in the diagnosis of deep endometriotic lesions and greater ability than other modalities to detect small lesions (Kinkel 2006). Recently, MRI was promoted as the non-invasive imaging technique of choice for detection and classification of endometriosis (Saba 2014a). Several recent systematic reviews on imaging in endometriosis (Hudelist 2011b; Medeiros 2014; Guerriero 2015) and narrative reviews on the topic primarily addressed diagnostic performance of US and/or MRI for deep endometriosis, predominantly with bowel involvement.

To improve diagnostic performance, variations in ultrasound techniques have been used, including transvaginal ultrasonography with bowel preparation (TVUS-BP) (Goncalves 2010), use of water contrast in the rectum (RWC-TVS) (Menada 2008a) or vagina (sonovaginography (SVG)) (Dessole 2003) and three-dimensional ultrasonography (3D-US) (Grasso 2010). Several modifications have been made to conventional MRI such as use of T1/T2-weighted (T1/T2-w) images, including addition of fat suppression with or without contrast enhancement. Three-dimensional MRI (3D-MRI) has also been evaluated as a single test for endometriosis, and 3.0T MRI has been developed using the 3.0T Magnetom system (in contrast to the widely used 1.5T system) with incorporation of T1/T2-w, fat-suppressed and 3D sequences (Hottat 2009; Manganaro 2013; Thomeer 2014). Computed tomography (CT)-based imaging (Biscaldi 2007), barium enema (Ribeiro 2008a) and other techniques have been explored as diagnostic tests for endometriosis. Improvements in imaging technology over time have positively affected the diagnostic ability of the same type of imaging test to detect endometriosis. Re-evaluation of diagnostic

test accuracy by Cochrane methods for a variety of imaging modalities is needed.

Clinical pathway

Women who present with symptoms of endometriosis (dysmenorrhoea, dyspareunia, chronic pelvic pain, difficulty conceiving) generally are investigated by a gynaecological examination and pelvic ultrasound scan to exclude other pathologies, in keeping with international guidelines (ACOG Committee on Gynecology 2010; SOGC 2010; Dunselman 2014). No other standard investigative tests are available, and MRI is used conservatively because of its cost. If women seek pain management rather than conception, empirical treatment with progestogens or the combined oral contraceptive pill is commonly started. Diagnostic laparoscopy is considered if empirical treatment fails, or if women decline or do not tolerate empirical treatment. In women who have difficulty conceiving, laparoscopy can be undertaken before fertility treatment is provided (particularly if severe pelvic pain or endometriomas are present) or after failed ART (assisted reproductive technology) treatment. Endometriosis can be diagnosed during fertility investigations in women who have minimal or no pain symptoms.

On average, a delay of six to 12 years is seen from onset of symptoms to definitive diagnosis at surgery (Matsuzaki 2006). Rapid referral to a gynaecologist with the ability to perform diagnostic surgery is associated with shorter time to diagnosis (Greene 2009). Collectively, young women, women in remote and rural locations and women of lower socioeconomic status have reduced access to surgery and are less likely to obtain prompt diagnosis and/or localisation of endometriosis.

Prior test(s)

Most women who present with symptoms suggestive of endometriosis undergo a full history and physical examination and a routine gynaecological ultrasound before the decision is made to perform diagnostic surgery. However, no consensus exists on whether ultrasound or any other test should be used routinely as part of a standardised approach.

Role of index test(s)

A new diagnostic test can fulfil one of three roles.

- Replacement: used to replace an existing test by providing greater or similar accuracy, along with other advantages.
- Triage: used as an initial step in a diagnostic pathway to identify women who need to undergo further testing with an existing test. Although ideally a triage test has high sensitivity and specificity, it may have lower sensitivity but higher specificity than the current test, or vice versa. The triage test does not aim to improve the diagnostic accuracy of the existing test but rather to reduce the number of individuals undergoing an unnecessary diagnostic test.
- Add-on: used in addition to an existing test to improve diagnostic performance (Bossuyt 2008).

Ideally, a diagnostic test is expected to correctly identify all women with a specific disease and to exclude all who do not have that disease, in other words, it should have sensitivity and specificity of 100%. High sensitivity indicates that a small number of women who have a negative test do have the disease (i.e. small number

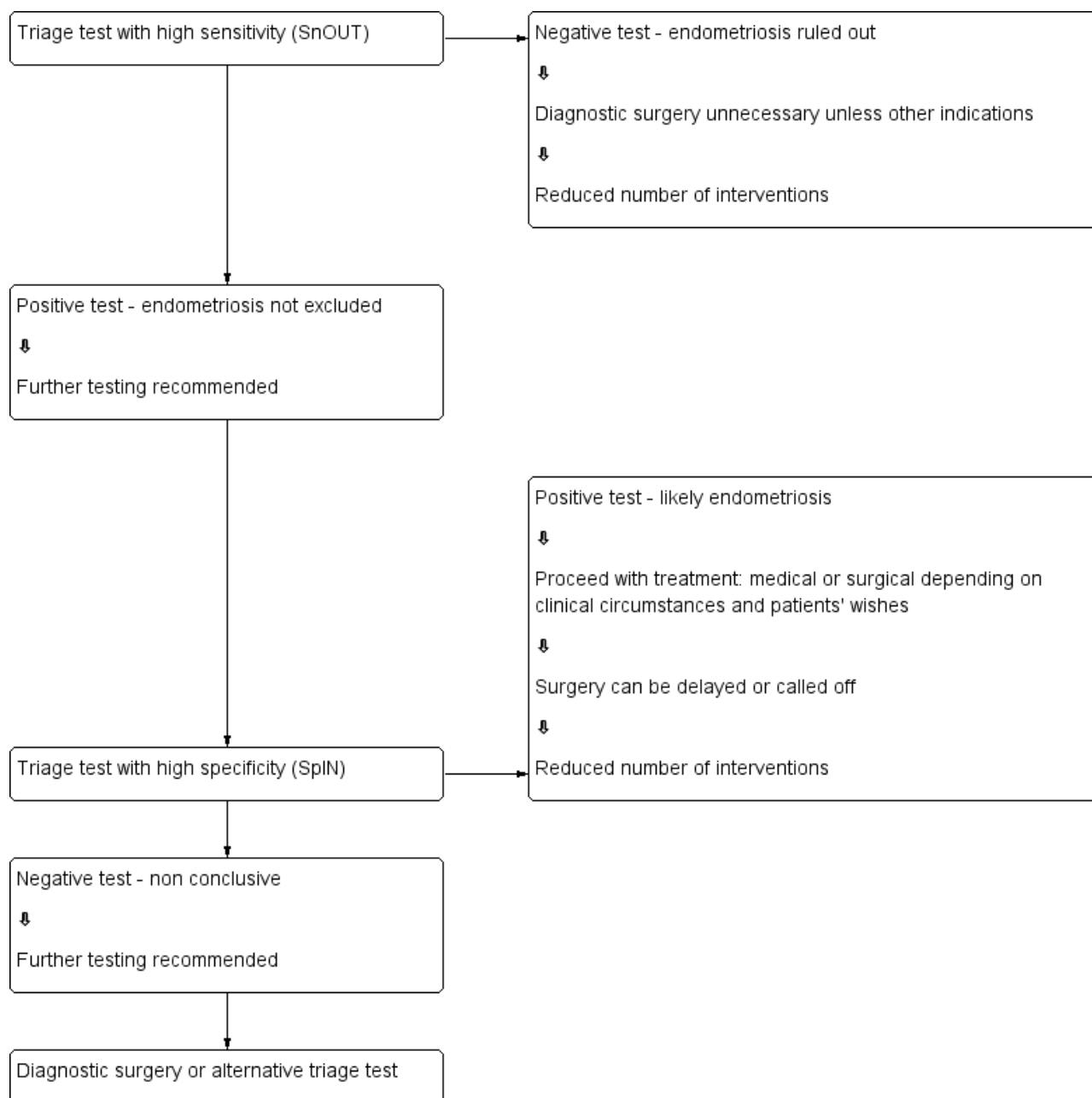
of false-negative results). High specificity corresponds to a small number of women who receive a positive test result but do not have the disease (i.e. small number of false-positive results). In practice, however, it is extremely rare to find a test with equally high sensitivity and specificity. An acceptable replacement test would need to have similar or higher sensitivity and specificity than the current gold standard of laparoscopy. The only systematic review performed to determine the accuracy of laparoscopy in diagnosing endometriosis reported sensitivity of 94% and specificity of 79% (Wykes 2004), and we have used this as a cut-off value for a replacement test.

The purpose of triage tests can vary depending on clinical context and patient priorities. One reasonable approach is to exclude the diagnosis to avoid further unnecessary and expensive diagnostic investigation. High-sensitivity tests yield few false-negative results and act to rule conditions out (SnNout). A negative result from a test with high sensitivity will exclude the disease with high certainty independent of the specificity. As women without disease would be assured of having a negative test, unnecessary invasive interventions can be avoided. However, a positive result has less diagnostic value, particularly when specificity is low. We predetermined that a clinically useful 'SnNout' triage test should have sensitivity of 95% or more and specificity of 50% or above. The sensitivity cut-off for a 'SnNout' triage test was set at 95% or above, if it is assumed that a 5% false-negative rate is statistically and clinically acceptable. The specificity cut-off was set at 50% or above, to avoid diagnostic uncertainty about more than 50% of the population receiving a positive result.

An alternative approach would be to avoid a missed diagnosis. High-specificity tests yield few false-positive results and act to rule conditions in (SpPin). A positive result for a highly specific triage test indicates a high likelihood of endometriosis. This information could be used to prioritise women for surgical treatment. A positive 'SpPin' test could also provide a clinical rationale for starting targeted disease-specific medical treatment for a woman without a surgical diagnosis, under the assumption that disease is present. Surgical management could be reserved for cases when conservative treatment fails. This is particularly relevant in some populations for which the therapeutic benefits of surgery for endometriosis have to be carefully balanced with the disadvantages (e.g. young women, women with medical conditions, pain-free women with a history of infertility). In this scenario, we considered sensitivity of 50% or above and specificity of 95% or higher as suitable cut-offs for a 'SpPin' triage test.

We evaluated imaging tests for their potential to replace surgery (replacement test) or to improve selection of women for surgery (triage test) that can rule out (SnNout) or rule in (SpPin) the disease. Both types of triage tests are clinically useful, minimising the number of unnecessary interventions. Sequential implementation of SnNout and SpPin tests can also optimise a diagnostic algorithm (Figure 1). We did not assess any test as an add-on test, as we sought tests that reduce the need for surgery - not tests that improve the accuracy of the currently available surgical diagnosis.

Figure 1. Sequential approach to non-invasive testing of endometriosis.



Knowledge of the accuracy of imaging index tests for detecting DIE at specific intrapelvic anatomical locations provides valuable information for surgeons, who can preoperatively arrange bowel preparation or availability of specialist surgical expertise for removal of lesions at particular locations. Surgical mapping of disease in isolated anatomical sites cannot exclude the disease somewhere else in the pelvis, hence it is not appropriate to use replacement test criteria for anatomical mapping, and we considered these types of tests only in the context of SnNout and SpPin triage criteria.

Alternative test(s)

No alternative tests for the diagnosis of endometriosis are available in routine clinical practice.

Rationale

Many women with endometriosis suffer long-standing pelvic pain and infertility before they receive the diagnosis. Surgery is the only method currently used to diagnose endometriosis, but it is associated with high costs and surgical risks. A simple and reliable non-invasive test for endometriosis with the potential to replace laparoscopy or to triage women to reduce surgery would minimise surgical risk and reduce diagnostic delay. Endometriosis could be detected at less advanced stages, and earlier interventions instituted. This would provide the opportunity for a preventative approach to this debilitating disease. Healthcare and social security costs of endometriosis would be expected to be reduced by early diagnosis and more cost-effective and efficient treatments.

Accurate diagnostic tests are important in strategic considerations of treatment planning. Women with severe invasive disease particularly benefit from surgical management, the efficacy of which depends on the completeness of excision of endometriotic lesions (Garry 1997). Therefore, ability to diagnose deep infiltrating endometriosis in general and at specific anatomical sites in particular might lead to selection of surgical technique, involvement of a multi-disciplinary surgical team or referral to the most appropriate practice (Chapron 2003a).

OBJECTIVES

Primary objectives

- To provide the estimates of the diagnostic accuracy of imaging modalities for the diagnosis of pelvic endometriosis, ovarian endometriosis and deeply infiltrating endometriosis (DIE) versus surgical diagnosis as a reference standard.
- To describe performance of imaging tests for mapping of deep endometriotic lesions in the pelvis at specific anatomical sites.

Imaging tests were evaluated as replacement tests for diagnostic surgery and as triage tests that would assist decision making regarding diagnostic surgery for endometriosis.

Secondary objectives

To investigate the influence of heterogeneity on the diagnostic accuracy of imaging modalities for endometriosis. Potential sources of heterogeneity include the following.

- Characteristics of the study population: age (adolescence vs later reproductive years); clinical presentation (subfertility, pelvic pain, ovarian mass, asymptomatic women); stage of disease (revised American Society for Reproductive Medicine (rASRM) classification system); geographic location of study.
- Histological confirmation in conjunction with laparoscopic visualisation versus laparoscopic visualisation alone.
- Changes in technology over time: year of publication; modifications applied to conventional imaging techniques.
- Methodological quality: differences in the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2) evaluation (low vs unclear or high risk); consecutive versus non-consecutive enrolment; blinding of surgeons to results of index tests.
- Study design ('single gate design' vs 'two-gate design' studies).

METHODS

Criteria for considering studies for this review

Types of studies

Published peer-reviewed studies that compared results of one or several types of imaging tests versus results obtained from a surgical diagnosis of endometriosis.

We included studies if they were:

- randomised controlled trials;

- observational studies of prospectively recruited women of the following designs:

- * 'single gate design' (studies with a single set of inclusion criteria defined by clinical presentation). All participants had clinically suspected endometriosis; or
- * 'two-gate design' (studies in which participants are sampled from distinct populations with respect to clinical presentation). The same study includes participants with a clinical suspicion of having the target condition (e.g. women with pelvic pain) and participants in whom the target condition is not suspected (e.g. women admitted for tubal ligation). Two-gate studies were eligible only when all cases and controls belonged to the same population with respect to the reference standard (i.e. all participants were scheduled for laparoscopy) (Rutjes 2005).

- performed in any healthcare setting; or
- published in any language;

We did not impose a minimal limit on the number of participants in the included studies nor on the number of studies that have evaluated each index test.

We excluded the following studies.

- Studies of specific study designs.
 - * Narrative or systematic review.
 - * Study of retrospective design when the index test was performed after execution of a reference test, or participants were selected through a retrospective review of case notes. Knowledge of the reference test could bias relatively subjective index tests. If endometriosis is found at a diagnostic surgical procedure, excision is commonly carried out concurrently, and this could bias the results of an index test performed after the reference standard.
 - * Case report or case series.
- Studies reported only in abstract form or in conference proceedings for which the full text was not available. This limitation was applied when we faced substantial difficulty in obtaining the information from abstracts, which precluded a reliable assessment of eligibility and methodological quality.

Participants

Study participants included women of reproductive age (puberty to menopause) with suspected endometriosis based on clinical symptoms and/or pelvic examination, who undertook both the index test and the reference standard.

Participants were selected from populations of women undergoing abdominal surgery for the following indications: (1) clinically suspected endometriosis (pelvic pain, infertility, abnormal pelvic examination or a combination of these), (2) ovarian mass regardless of symptoms, (3) a mixed group, which consists of women with suspected endometriosis/ovarian mass and/or women with other benign gynaecological conditions (e.g. surgical sterilisation, fibroid uterus).

Articles that included participants of postmenopausal age were eligible when data for the reproductive age group were available in isolation. Studies were excluded when the study population involved participants who clearly would not undergo the index test in a clinical scenario and/or would not benefit from the test (e.g.

women with ectopic pregnancies, gynaecological malignancies, acute pelvic inflammatory disease). We also excluded publications in which only a subset of participants with a positive index test or reference standard were included in the analysis, and data for the whole cohort were not available.

Index tests

All types of imaging modalities for endometriosis, including possible modifications to conventional techniques, were assessed separately or in combination with other imaging tests. We attempted to group several types of tests that were based on common technical principles and similarity in clinical applicability. The index tests assessed are presented and described in [Table 2](#).

We considered studies only if data were reported in sufficient detail for construction of 2×2 contingency tables. We included only studies that reported diagnostic accuracy estimates per number of participants ('participant-level' analysis).

We undertook an independent evaluation of the diagnostic test accuracy of imaging tests to anatomically map endometriotic lesions because multiple endometriotic implants can co-exist at different sites in the same individual. For this 'region level' analysis, only analyses that recorded data estimates per number of participants were included, as information about the accuracy of imaging tests for mapping the disease is more informative and clinically applicable when presented as per-participant calculations of accuracy estimates.

Combined evaluations of imaging tests and other methods of diagnosing endometriosis (e.g. pelvic examination; urine, endometrial or blood tests) are beyond the scope of this review and are presented separately in another review titled 'Combined tests for the non-invasive diagnosis of endometriosis'. We excluded from the review studies that solely assessed specific technical aspects, radiological criteria or interobserver variability of index tests without reporting data on diagnostic performance.

The diagnostic performance of an index test was considered high when the test reached the criteria for a replacement test (sensitivity $\geq 94\%$ with specificity $\geq 79\%$) or a triage test (sensitivity $\geq 95\%$ with specificity $\geq 50\%$, or vice versa). We categorised as 'approaching' high accuracy imaging tests with diagnostic estimates within 5% of set thresholds. We considered all other diagnostic estimates as low.

Target conditions

Investigators assessed three target conditions.

- Pelvic endometriosis: defined as endometrial tissue located within the pelvic cavity, including any of the pelvic organs, peritoneum and pouch of Douglas.
- Ovarian endometriosis (endometrioma): defined as ovarian cysts lined by endometrial tissue and appearing as an ovarian mass of varying size.
- DIE: defined as subperitoneal infiltration of endometrial implants, for example, when endometriotic implants penetrate the retroperitoneal space for a distance of 5 mm or more. Posterior DIE is the most common form of DIE, and both conditions are interchangeably reported. For the purpose of this review, we combined them as a single target condition - DIE/posterior DIE.

In addition, the ability of diagnostic imaging to map endometriotic lesions at specific anatomical pelvic locations was evaluated. Anatomical locations included rectovaginal septum (RVS), uterosacral ligament (USL), vaginal wall, POD obliteration, anterior DIE, rectosigmoid colon and the entire bowel from ileum to rectum. These locations are defined in [Table 3](#).

Certain rare types of endometriosis such as extrapelvic, bladder and ureteric endometriosis were not included in this review because most of these were described in case reports or in case series, and laparoscopy and laparotomy are not reliable reference standards for these conditions.

We excluded studies in which the diagnosis of endometriosis was not the primary outcome of the trial (e.g. malignant vs benign masses, normal vs abnormal pelvis) and separate data for endometriosis were not available.

We also excluded studies in which findings of the index test formed the basis of selection for the reference standard because this was likely to distort any assessment of the diagnostic value of the index test.

We included studies that involved only selected populations of women with endometriosis (i.e. at specific rASRM stages), in view of emerging evidence on poor correlation of this classification with infertility and pain symptoms. Exclusion of these studies could result in loss of potentially important diagnostic information from otherwise eligible publications. When possible, we addressed the impact of these studies in investigations of heterogeneity. When a study analysed a large population with a wide spectrum of endometriosis and additionally reported subgroup analyses of different stages of disease severity, we considered only estimates for the entire population because subgroup analyses do not directly address the review question regarding clinical utility of biomarkers in detecting the disease.

Reference standards

The reference standard was visualisation of endometriosis at surgery (laparoscopy or laparotomy) with or without histological confirmation, as this is currently the best available test for endometriosis. We reviewed information regarding interobserver and intraobserver correlation of the reference standard, if reported.

We included only studies in which the reference test was performed within 12 months of the index test, on the assumption that disease status could change within a period of one year or longer, naturally or as a result of treatment. We did not include in this review studies in which the participants did not undergo the reference standard, or for whom findings of the index test formed the basis of selection for the reference standard.

Summary of inclusion/exclusion criteria

Inclusion criteria

- Types of studies.
 - * Published peer-reviewed.
 - * Randomised controlled trials (RCTs).
 - * Observational with prospective recruitment in the following design.
 - ☐ 'Single-gate design' (single set of inclusion criteria defined by clinical presentation) - all participants had clinically suspected endometriosis.
 - ☐ 'Two-gate design' (two sets of inclusion criteria with respect to clinical presentation and one set of inclusion criteria with respect to reference standard) - participants with or without clinical suspicion of endometriosis scheduled for abdominal surgery.
 - * Published in any language.
 - * Performed in any healthcare setting.
 - * Any sample size.
- Participants.
 - * Women of reproductive age.
 - * Women with clinically suspected endometriosis, including women who underwent abdominal surgery for other benign gynaecological conditions and surgical assessment for presence/absence of endometriosis.
 - * Those who undertook both index test and reference standard.
- Index tests.
 - * One or several types of imaging tests.
 - * Data reported in sufficient detail for construction of 2 × 2 tables and presented as 'participant-level' analysis.
- Target conditions.
 - * Pelvic endometriosis, ovarian endometrioma, DIE or specific pelvic sites of DIE.
- Reference standard.
 - * Surgical visualisation of lesions for the diagnosis of endometriosis (laparoscopy or laparotomy) with or without histological verification.
 - * Performed within 12 months of the index test.

Exclusion criteria

- Types of studies.
 - * Narrative or systematic reviews.
 - * Retrospective design in which the index test was performed after execution of the reference test and/or participants were selected from a retrospective review of case notes.
 - * Case reports or case series.
 - * Conference proceedings.
- Participants.
 - * Included cohort was not representative of the target population that would benefit from the test (e.g. women with known genital tract malignancy, ectopic pregnancy, acute pelvic inflammatory disease).
 - * Study included participants of postmenopausal age, and data for the reproductive age group were not available in isolation.
 - * Only participants with positive index test or positive reference standard were included in the analysis.

- Index tests.
 - * Imaging tests were presented in combination with other diagnostic tests for endometriosis, and separate information was not available for the imaging modalities.
 - * Study presented only specific technical aspects of an index test or data on interobserver variability, rather than diagnostic performance of the test.
 - * Study presented only qualitative description of radiological appearance of endometriotic lesions.
 - * Only the number of lesions rather than the number of participants with endometriosis was reported (i.e. 'lesion-level' analyses).
- Target conditions.
 - * Endometriosis was not the primary outcome of the trial (e.g. malignant vs benign masses, normal vs abnormal pelvis).
 - * Atypical, rare sites of endometriosis.
- Reference standard.
 - * Reference standard performed only in a subset of study/control group.
 - * Findings of the index test formed the basis of selection for the reference standard.
 - * Other than specified in inclusion criteria.

Search methods for identification of studies

The search strategy was developed in collaboration with the Trials Search Co-ordinator of the Gynaecology and Fertility Review Group, according to recommendations provided in *the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (de Vet 2008). Searches were not limited to particular types of study design and did not apply language or publication date restrictions. The search strategy incorporated words in the title, abstracts, text words across the record and medical subject headings (MeSH). The search was created initially for one broad review examining all diagnostic tests for endometriosis, but because of the complexity of this review, it was split into five separate reviews, and separate searches were used for imaging tests and for biomarker tests (endometrial, blood, urinary, combined). All searches were performed from inception until present. Search strategies for each database and the number of hits per search are presented in [Appendix 1](#). A summary of search results is presented under [Results of the search](#).

Electronic searches

We searched the following databases to identify published articles that assessed the diagnostic value of imaging tests for endometriosis.

- MEDLINE.
- EMBASE.
- Cochrane Central Register of Controlled Trials (CENTRAL).
- Cumulative Index to Nursing and Allied Health Literature (CINAHL).
- PsycINFO.
- Web of Science Core Collection.
- Latin American Caribbean Health Sciences Literature (LILACS).
- Open Archives Initiative database (OAIster).
- Turning Research Into Practice database (TRIP).

- Databases of trial registers.
 - * ClinicalTrials.gov.
 - * World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) search portal.
- Databases to identify reviews as a source of references to potentially relevant studies.
 - * Database of diagnostic studies and diagnostic reviews (MEDION).
 - * Database of Abstracts of Reviews of Effects (DARE).
 - * PubMed, a 'Systematic Review' search under the 'Clinical Queries' link.
- Searches for papers recently published and not yet indexed in the major databases.
 - * PubMed (simple search of the past six months).

Searching other resources

We handsearched reference lists of all relevant publications (retrieved full texts of key articles and identified reviews).

We abandoned an attempt to locate grey literature (unpublished studies and conference proceedings), as we faced substantial difficulty in obtaining full-text publications or additional details of studies reported in abstract form. This precluded reliable assessment of eligibility and methodological quality of studies, and it was decided that we would not include these publication sources in this review.

Data collection and analysis

Selection of studies

One review author (VN) scanned the titles of studies identified by our search to remove clearly irrelevant articles, and reviewed titles and abstracts of remaining studies to select potentially relevant publications. Two review authors (VN and LH) independently reviewed full-text versions of articles selected by title and abstract, and assessed eligibility for inclusion on the basis of criteria listed above under [Criteria for considering studies for this review](#). A single failed eligibility criterion was sufficient for a study to be excluded from the review.

Review authors who assessed the relevance of studies and eligibility for inclusion were not blind to information about each article, including publishing journal, names of authors, institutions and results. Disagreements were resolved by discussion and, if necessary, by consultation with a third review author (CF), who is an expert in the field and in methodological aspects of Cochrane systematic reviews.

When papers updated previous publications and were performed on the same study population at different recruitment points, we used the most complete data set that superseded previous publications to avoid double counting of participants or studies. We retrieved missing data by directly contacting the authors to clarify study eligibility. When potentially relevant studies were found in languages other than English, we arranged for a translation. For excluded studies, we documented reasons for exclusion and details of which criteria were not met. We have presented characteristics of included, excluded and awaiting classification studies under [Characteristics of included studies](#), [Characteristics of excluded studies](#) and [Characteristics of studies awaiting classification](#), respectively.

Data extraction and management

Two independent review authors (VN, LH) extracted data from eligible studies and resolved disagreements by consulting with a third review author (CF). If required, we contacted study investigators to resolve questions regarding the data.

To collect details from included studies, we specifically designed for this review a data extraction form and pilot-tested it on three studies of diagnostic accuracy tests for endometriosis. We recorded the following information for each study.

- General information and study design: first author, year of publication, country, language, setting, objectives, inclusion/exclusion criteria, type of enrolment.
- Characteristics of study participants: age, symptoms/history/previous tests, type of target condition and its prevalence in the study population, number of participants enrolled and available for analysis, reasons for withdrawal.
- Features of the index test and the reference standard: type, diagnostic criteria, number and experience of the operators, blinding of operators to other tests and/or clinical data, interobserver variability, time interval between index test and reference standard.
- Reported numbers of true-positives (TPs), false-negatives (FNs), true-negatives (TNs) and false-positives (FPs) were used to construct a 2×2 table for each index test. If these values were not reported, we attempted to reconstruct 2×2 tables from the diagnostic estimates presented in the article.

We extracted data into Review Manager (RevMan) software, which was used to graphically display quality assessment, diagnostic estimates data and descriptive analyses.

Assessment of methodological quality

We used QUADAS-2, a modified version of the QUADAS tool, to assess the quality of each included study ([Whiting 2011](#)).

We have presented the review-specific QUADAS-2 tool and an explanatory document in [Table 4](#). We judged each paper as having a 'low', 'high' or 'unclear' risk of bias for each of four domains, and we assessed concerns about applicability in three domains. We considered studies as having low methodological quality when classified at high or unclear risk of bias and/or high concern regarding applicability in at least one domain. Two review authors (VN, LH) independently assessed each included study and settled disagreements by reaching consensus. Two review authors independently piloted the topic-specific tool to rate four of the included studies at a high level of agreement. We made the following modifications (specific to the imaging modalities review) to signalling questions of the original QUADAS-2 tool.

Domain 1

- An original signalling question 'Was a case control design avoided?' was rephrased as 'Was a two-gate design avoided?'. Diagnostic studies are cross-sectional in nature, aiming to compare results of an index test versus results of the reference standard in the same group of participants. In these studies, parameters are measured at a single point in time, and groups are classified by the outcome of the reference standard, albeit the analysis is performed retrospectively. Therefore, unlike in epidemiological studies, the terms 'cohort' and 'case-

control' are less informative for diagnostic test trials and were substituted by 'single-gate' and 'two-gate' designs. We included this question because a two-gate design has greater potential to introduce selection bias.

Domain 2

- An additional signalling question 'Was the index test performed by a single operator?' was included to assess interobserver variation bias.
- An additional signalling question 'Were the same clinical data available when the index test results were interpreted as that which would be available when the test is used in practice?' was included to assess bias in clinical applicability.
- An original signalling question 'If a threshold was used, was it prespecified?' was rephrased as 'Did the study provide a clear prespecified definition of what was considered to be a positive index test result?' because this question was more applicable to imaging modalities.

We assessed methodological quality for each domain but did not calculate a summary score to estimate the overall quality of studies (Whiting 2005).

Statistical analysis and data synthesis

We analysed diagnostic imaging techniques in the following subsets.

- Tests for detecting pelvic endometriosis.
- Tests for detecting ovarian endometriosis (endometrioma).
- Tests for detecting DIE.
- Tests for identifying deep endometriotic lesions at separate pelvic anatomical sites (USL, RVS, vaginal wall, obliterated POD, rectosigmoid colon, bowel (ileum to caecum)).

We generated estimates of sensitivity and specificity in forest plots and plotted them in the receiver operating characteristic (ROC) space for each index test using RevMan. We investigated the diagnostic performance of each test and visually explored interstudy variation in performance of each index test in relation to participant characteristics, study design and study quality factors. We included two or more tests evaluated in the same cohort as separate data sets because the unit of analysis was the test result - not the participant.

We obtained the estimate of the expected operating point (mean sensitivity and specificity) and corresponding 95% confidence region by using the bivariate logit normal random-effects model for all meta-analyses including four or more studies. When fewer than four studies were included, we did not attempt to estimate co-variance and reported this as zero. To estimate the performance of other tests in small meta-analyses (two to three data sets), we performed a fixed-effect meta-analysis in the absence of substantial heterogeneity, resulting in the summary estimate for sensitivity and for specificity. We performed meta-analyses by using SAS NLMIXED software. We entered results from SAS into RevMan to provide plots of estimated mean or summary points and confidence regions, superimposed on study-specific estimates of sensitivity and specificity.

We assessed the comparative accuracy of index tests for each target condition in two ways. In direct, fully paired comparisons in which

all study participants received more than one index test, as well as the reference standard, we plotted the estimates in RevMan. If meta-analysis was possible, we used test-level co-variables in the bivariate logit normal model to identify statistically significant differences. Otherwise, we reported available comparative data in a narrative way and illustrated the data using forest and ROC plots.

When test performance was judged against predetermined diagnostic criteria, we considered the point estimates of sensitivity and specificity as the most informative presentation of test performance. We acknowledge that tests with point estimates that did not reach the predetermined criteria but included confidence intervals (CIs) that contained values above the threshold could have provided diagnostic value. Furthermore, tests with point estimates that reached the criteria but with CIs that contained values below the threshold could have provided overestimated diagnostic value. If the range of CIs rather than the point estimates of data are used, the predetermined cut-off becomes meaningless. Therefore we did not consider CIs in qualifying the test performance but used this information when interpreting reliability of the data obtained.

Dealing with missing data

We defined missing data as any information regarding study population, index tests or reference standards that was not available in the publication but was required to determine the eligibility of the study for inclusion, to assess its methodological quality or to construct results tables. If we identified missing data, we contacted study authors in an attempt to obtain this information. If missing data prevented a clear judgement regarding applicability for inclusion or construction of accurate 2 × 2 tables, and if data were not provided by the primary investigators (e.g. we were not able to locate contact details of study authors, we received no reply from study authors, study authors replied that the requested information was unavailable), we excluded the study from the review.

Investigations of heterogeneity

We initially assessed heterogeneity by visually examining forest plots of sensitivities and specificities and ROC plots for all index tests. We stated potential sources of heterogeneity under [Secondary objectives](#). For diagnostic tests that involved more than 10 eligible studies or data sets, we planned to formally explore heterogeneity by using study level co-variables. We also planned to assess the sensitivity of results to inclusion and exclusion of outlying studies in all analyses but refrained from doing so because of the small number of studies available for most analyses. It is important to use caution when interpreting small meta-analyses (few studies) with a limited total sample size.

Sensitivity analyses

We planned to conduct sensitivity analyses to assess the impact of the methodological quality of included studies on results of the meta-analysis, if sufficient data were available. We defined low-quality studies as having high risk of bias for one or more QUADAS-2 domains. We also planned to use the 'leave-one-out' procedure to assess the impact of each study on results of the meta-analysis (leading study effect), but we were not able to do this because of the small number of studies available for most groups of tests.

Assessment of reporting bias

A comprehensive search of multiple sources for eligible studies, a search of trial registers and application of no language restrictions minimised the risk of reporting bias. However, publication bias generally arises when studies have a greater chance of being published if their results are positive. Therefore, we initially searched unpublished and published study databases and conference proceedings and evaluated identified sources. During the process of qualifying studies for inclusion in this review, we faced substantial difficulty in obtaining full-text publications or additional details of studies published in abstract form. This precluded reliable assessment of eligibility and methodological quality, and it was decided to excluded these publication sources from this review.

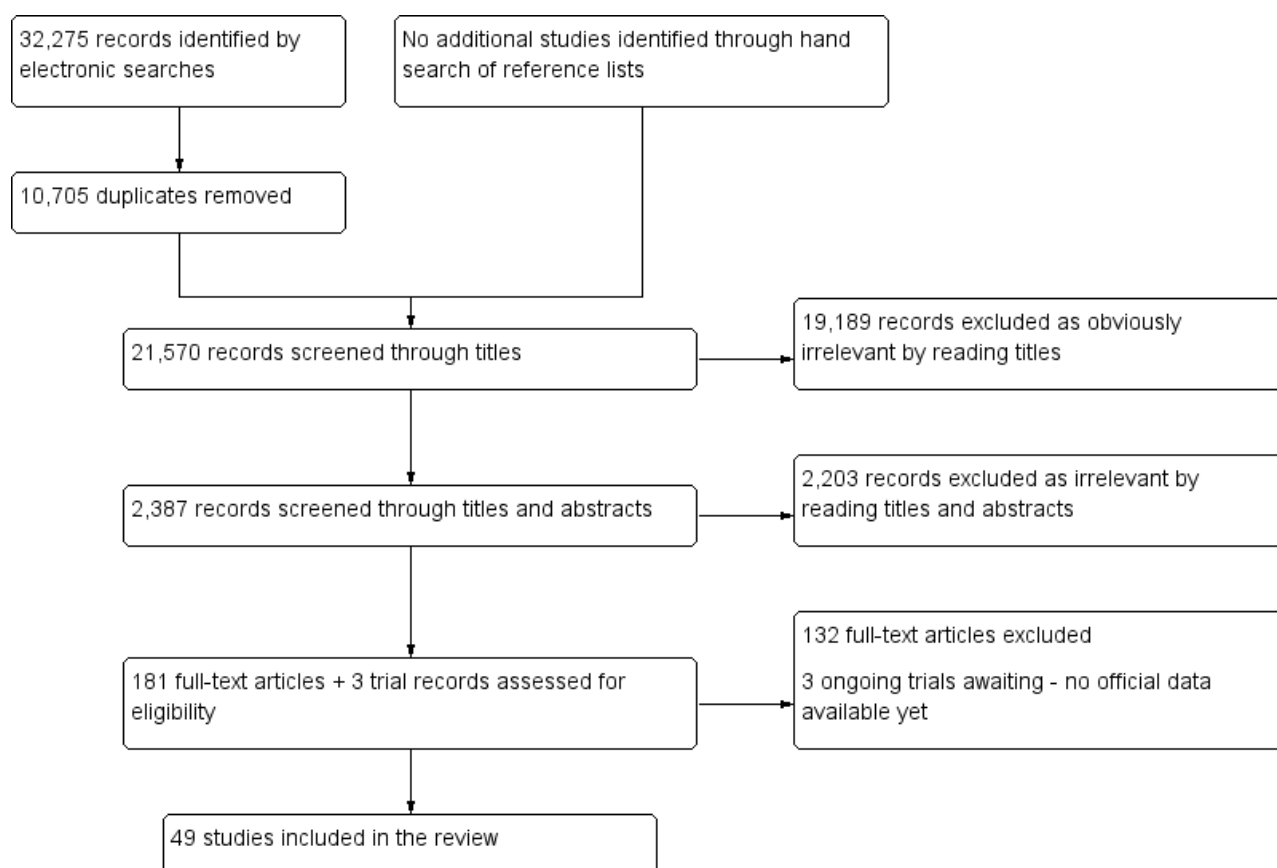
RESULTS

Results of the search

The literature search identified 32,275 references as follows: MEDLINE (n = 7391), EMBASE (n = 12,161), CENTRAL (n = 445),

CINAHL (n = 668), PsycINFO (n = 174), Web of Science (n = 7425), LILACS (n = 420), OAlster (n = 446), TRIP (n = 1648), trial registers for ongoing and registered trials (n = 523), MEDION (n = 190), DARE (n = 99), PubMed, a 'Systematic Review' search (n = 418) and simple search PubMed (n = 267). We searched these databases from inception to 20 April. The flow of the selection process is presented in [Figure 2](#). We screened titles to exclude duplicates (n = 10,705) and clearly irrelevant studies (n = 19,189). We eliminated another 2205 references eliminated after reading the abstracts because they did not address the research question or clearly did not meet the inclusion criteria. We retrieved the full texts of the remaining 181 references and assessed them for eligibility. Data from 63 studies required additional clarification from study authors and 25 non-English publications were translated. Ultimately, 49 studies that were eligible according to the inclusion criteria provided data for the review; we excluded 132 studies. In addition, we identified three ongoing trials through the clinical trials registries ([Characteristics of ongoing studies](#)) but found that the outcomes of these studies were not yet available (two trials were still open to participant recruitment, and the status of one study was unclear). We will monitor and address the progress of these studies in future updates.

Figure 2. Flow of studies identified in literature search for systematic review on imaging modalities for a non-invasive diagnosis of endometriosis.



Basic features of included studies

We have presented the list and details of the included studies under [Characteristics of included studies](#). The 49 included studies studied 4807 participants, with a median of 87 women per study (range

10 to 710). Of these studies, 27 were conducted in Europe, six in South America, four in Asia, two in North America, three in Australia and one in the Middle East. Ninety-four per cent (46/49) of these studies were conducted at university hospitals, of which 14 were designated as referral centres for endometriosis. The earliest article

was published in 1993, 41 articles were published after 2000 and 26 after 2010. All included studies assessed women of reproductive age; 46 studies included a population with clinical suspicion of endometriosis based on symptoms and clinical examination with or without an ovarian mass. Two studies assessed only women with a persistent ovarian mass (Guerriero 1996a; Guerriero 1996b), and one study focused exclusively on women undergoing infertility workup (Ubaldi 1998). Only one study (Eskenazi 2001) used a 'two-gate design' and included a wider group of participants, defined as 'women scheduled to undergo laparoscopy or laparotomy for pelvic pain, infertility, tubal ligation, or masses of the adnexa or uterus'. Eleven studies (Okada 1995; Guerriero 1996a; Guerriero 1996b; Ghezzi 2005; Takeuchi 2005; Chamie 2009a; Fastrez 2011; Manganaro 2012a; Manganaro 2012b; Manganaro 2013; Mangler 2013) reported abnormal imaging findings (other than the index test) as one of the inclusion criteria, but the remaining studies presented no information on pre-enrolment imaging tests. One study limited the study population to 'women with symptoms suggestive of endometriosis with normal ovarian size and no evidence of an ovarian cyst' (Said 2014). Seventeen studies (Stratton 2003; Biscaldi 2007; Bazot 2009; Hottat 2009; Piketty 2009; Bergamini 2010; Falco 2011; Fastrez 2011; Ferrero 2011; Savelli 2011; Mangler 2013; Reid 2013a; Stabile 2013; Biscaldi 2014; Guerriero 2014; Piessens 2014; Said 2014) included women with a history of previous surgery for endometriosis representing 7% to 66% of the study population. Two studies (Holland 2010; Mangler 2013) included participants with a recent laparoscopic diagnosis of endometriosis who were awaiting definitive surgery; however, index test operators were blind to previous surgical findings. Nine studies described exclusion of participants who had undergone any pelvic surgery (Dessole 2003; Ghezzi 2005; Takeuchi 2005; Chamie 2009a; Biscaldi 2014; Said 2014) or specific excision of DIE (Fedele 1998; Hudelist 2011a; Hudelist 2013). Laparoscopy was the predominant surgical modality in all studies, whereas laparotomy was reserved for selected cases. Eighty-eight per cent (43/49) of the included studies used histopathology to confirm the surgical diagnosis. The reported prevalence of endometriosis varied, ranging from 43% to 100% for pelvic endometriosis, from 7.5% to 100% for ovarian endometriosis and from 30% to 100% for DIE.

Authors of five papers declared that they received no financial support from external sources (Ribeiro 2008a; Hottat 2009; Fastrez 2011; Manganaro 2012b; Manganaro 2013). Guerriero 2014 stated that this study was partially supported by the Regione Autonoma della Sardegna (project code CPR-24750) but declared no conflict of interest. Stratton 2003 and Hudelist 2013 declared that work was funded by the Intramural Program, National Institute of Child Health and Human Development, Bethesda, Maryland, and by

the OEGEO, Österreichische Gesellschaft für Endokrinologische Onkologie, respectively, but made no statement regarding a conflict of interest. Nine other articles declared no conflict of interest (Guerriero 2008; Bazot 2009; Hottat 2009; Fastrez 2011; Manganaro 2012b; Manganaro 2013; Mangler 2013; Said 2014; Thomeer 2014), and the remaining included studies provided no such information.

Basic features of excluded studies

We have presented the list and descriptions of excluded studies under [Characteristics of excluded studies](#). On the basis of full-text assessment, we excluded 132 publications, 34 of which were of retrospective design by which the population was selected from medical records, and index tests were reviewed retrospectively. We excluded an additional 26 studies as they reported outcomes for number of lesions - not number of participants (a 'lesion-level' analysis). Twenty-six studies were not diagnostic test accuracy studies and focused on technical aspects of the test, interobserver variability or a description of radiological criteria of the target condition. We excluded 11 papers as they enrolled a wide age group ($n = 9$) or pregnant women ($n = 2$), and independent assessment of women of reproductive age could not be performed. Many articles in this excluded group were comparisons of endometriomas versus benign and malignant ovarian masses in older women. In eight excluded papers, a reference standard other than surgery was used, or investigators provided no data on surgical diagnosis. In nine of the excluded studies, the target condition was outside the inclusion criteria, and data were reported for benign versus malignant masses or normal versus abnormal pelvis with no separate data given for endometriosis. We excluded another eight studies as they reported on a cohort that overlapped with a cohort in another updated included paper. Four studies presented insufficient descriptions of methods and/or study populations and provided no information. We could not extract data for 2×2 tables from three studies. For two other studies, the index test was outside the inclusion criteria, reporting data for a combination of imaging tests and pelvic examination. We excluded one study as investigators did not consider healthy controls in the analysis, and another study because the time interval between index test and surgery exceeded 12 months.

Methodological quality of included studies

We have presented details on the quality of included studies in the QUADAS-2 results summary ([Figure 3](#) and [Figure 4](#)). Overall, most studies were of poor methodological quality, and only one study (Thomeer 2014) was assigned low risk of bias in every domain assessed.

Figure 3. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.

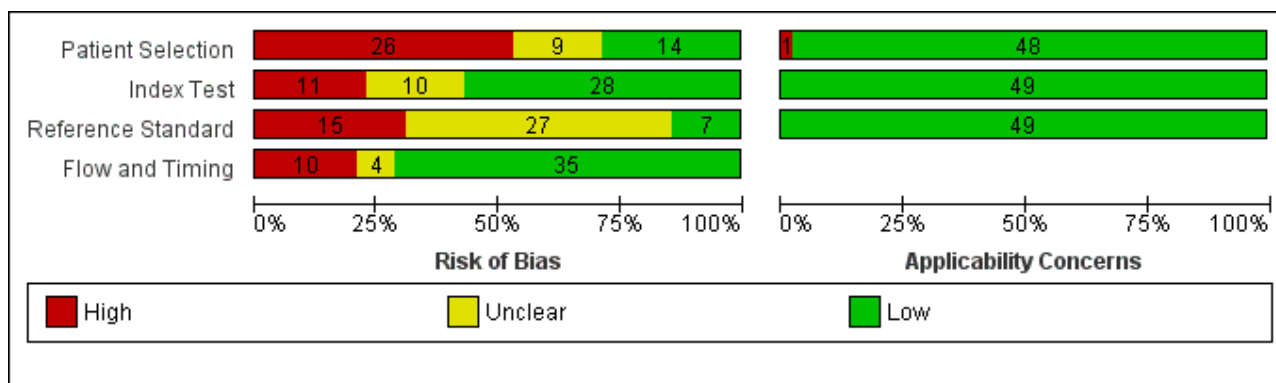


Figure 4. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Abrao 2007	+	+	-	+	+	+	+
Ascher 1995	-	+	?	-	+	+	+
Bazot 2009	?	+	?	+	+	+	+
Bazot 2013	?	+	?	-	+	+	+
Bergamini 2010	-	-	-	?	+	+	+
Biscaldi 2007	-	+	?	+	+	+	+
Biscaldi 2014	-	+	-	+	+	+	+
Chamie 2009a	-	+	?	?	+	+	+
Dessole 2003	-	-	-	?	+	+	+
Eskenazi 2001	-	+	?	+	-	+	+
Falco 2011	-	+	?	-	+	+	+
Fastrez 2011	-	-	+	+	+	+	+
Fedele 1998	-	-	-	+	+	+	+
Ferrero 2011	-	-	-	+	+	+	+
Ghezzi 2005	+	?	?	+	+	+	+
Goncalves 2010	+	+	?	+	+	+	+
Grasso 2010	-	+	?	+	+	+	+
Guerriero 1996a	+	?	?	+	+	+	+
Guerriero 1996b	+	?	?	+	+	+	+
Guerriero 2007	?	?	?	+	+	+	+
Guerriero 2008	?	?	?	+	+	+	+
Guerriero 2014	+	+	-	-	+	+	+
Ha 1994	-	+	?	+	+	+	+
Holland 2010	+	-	+	+	+	+	+
Hottat 2009	+	+	+	-	+	+	+
Hudelist 2011a	-	+	-	+	+	+	+

Figure 4. (Continued)

Hudelist 2011a	+	+	+	+	+	+	+
Hudelist 2013	+	+	+	+	+	+	+
Leon 2014	+	?	+	+	+	+	+
Manganaro 2012a	+	?	?	+	+	+	+
Manganaro 2012b	+	+	?	+	+	+	+
Manganaro 2013	+	+	?	+	+	+	+
Mangler 2013	?	+	+	+	+	+	+
Menada 2008a	+	+	?	+	+	+	+
Okada 1995	+	?	?	+	+	+	+
Pascual 2010	?	?	?	+	+	+	+
Piessens 2014	?	+	+	+	+	+	+
Piketty 2009	+	+	?	+	+	+	+
Reid 2013a	+	+	?	+	+	+	+
Reid 2014	+	+	+	+	+	+	+
Ribeiro 2008a	+	+	+	+	+	+	+
Said 2014	+	+	?	+	+	+	+
Savelli 2011	+	+	+	+	+	+	+
Scarella 2013	+	?	+	+	+	+	+
Stabile 2013	+	+	?	+	+	+	+
Stratton 2003	?	+	+	+	+	+	+
Sugimura 1993	?	+	?	+	+	+	+
Takeuchi 2005	+	+	+	?	+	+	+
Thomeer 2014	+	+	+	+	+	+	+
Ubaldi 1998	+	+	?	+	+	+	+

+ High
 ? Unclear
 + Low

Twenty-six studies presented high risk of patient selection bias (Ha 1994; Ascher 1995; Okada 1995; Fedele 1998; Ubaldi 1998; Eskenazi 2001; Dessole 2003; Takeuchi 2005; Biscaldi 2007; Menada 2008a; Ribeiro 2008a; Chamie 2009a; Piketty 2009; Bergamini 2010; Grasso 2010; Falco 2011; Fastrez 2011; Ferrero 2011; Hudelist 2011a; Manganaro 2012a; Manganaro 2012b; Manganaro 2013; Reid 2013a; Biscaldi 2014; Leon 2014; Said 2014), nine were rated as having unclear risk (Sugimura 1993; Stratton 2003; Guerriero 2007; Guerriero 2008; Bazot 2009; Pascual 2010; Bazot 2013; Mangler 2013; Piessens 2014) and 14 demonstrated low

risk. Non-consecutive or non-random enrolment, absence of a clear definition of inclusion/exclusion criteria and inclusion of a highly selected group of participants were the main reasons for assessment of high risk of bias.

Eleven studies presented with high risk of index test interpretation bias (Sugimura 1993; Fedele 1998; Dessole 2003; Bergamini 2010; Holland 2010; Fastrez 2011; Ferrero 2011; Savelli 2011; Mangler 2013; Piessens 2014; Reid 2014), 10 demonstrated unclear risk (Okada 1995; Guerriero 1996a; Guerriero 1996b; Ghezzi 2005;

Guerriero 2007; Guerriero 2008; Pascual 2010; Manganaro 2012a; Scarella 2013; Leon 2014) and 28 were rated as having low risk. Lack of clear prespecified criteria for a positive diagnosis and lack of blinding of index test operators to the clinical history or to results of other diagnostic tests were the main reasons for high risk assessment. High risk of bias for this domain was also attributed to articles in which the index test was performed/interpreted by different operators for different participants, as varying skill levels could undermine the reliability of results. Overall, interobserver variability was rarely reported. Six studies stated that disagreement between test operators was resolved by consensus in a joint session (Ascher 1995; Ghezzi 2005; Biscaldi 2007; Chamie 2009a; Manganaro 2012a; Thomeer 2014); two calculated accuracy estimates of the index test separately for the two examiners (Hottat 2009; Holland 2010) and eight assessed interobserver and intraobserver variability in the whole cohort or in a subset of randomly selected participants (Ubaldi 1998; Guerriero 2008; Hottat 2009; Manganaro 2012b; Bazot 2013; Stabile 2013; Guerriero 2014; Thomeer 2014). None of the included studies carried risk of test review bias, as studies in which the index test was performed or interpreted after execution of the reference standard were excluded. As criteria for a positive index test were variable between studies and as index test protocols were not standardised, quality judgements for the index test were complex; however, these factors were not directly addressed by the QUADAS-2 tool.

Fifteen studies were at high risk of bias in the 'reference standard' domain (Fedele 1998; Dessole 2003; Abrao 2007; Ribeiro 2008a; Bergamini 2010; Ferrero 2011; Hudelist 2011a; Savelli 2011; Hudelist 2013; Mangler 2013; Biscaldi 2014; Guerriero 2014; Leon 2014; Piessens 2014; Reid 2014), 27 were classified as unclear risk (Sugimura 1993; Ha 1994; Ascher 1995; Okada 1995; Guerriero 1996a; Guerriero 1996b; Ubaldi 1998; Eskenazi 2001; Ghezzi 2005; Biscaldi 2007; Guerriero 2007; Guerriero 2008; Menada 2008a; Bazot 2009; Chamie 2009a; Piketty 2009; Goncalves 2010; Grasso 2010; Pascual 2010; Falco 2011; Manganaro 2012a; Manganaro 2012b; Bazot 2013; Manganaro 2013; Reid 2013a; Stabile 2013; Said 2014) and seven demonstrated low risk. We assigned high risk of bias when reference standards were interpreted with knowledge of index test results. Although it would be unethical to withhold from surgeons information on preoperative imaging investigations, lack of blinding to the index test contributes to diagnostic review bias. Most studies provided insufficient information to indicate how likely the reference standard was to have correctly classified the target condition. Specifically, surgical procedures were not well described, criteria for a positive reference standard were not stated or no information was provided regarding the experience of the surgeons and/or pathologists involved.

Ten studies were at high risk of bias in the 'flow and timing' domain (Ascher 1995; Stratton 2003; Hottat 2009; Falco 2011; Savelli 2011; Bazot 2013; Scarella 2013; Guerriero 2014; Leon 2014; Piessens 2014), four were at unclear risk (Dessole 2003; Takeuchi 2005; Chamie 2009a; Bergamini 2010) and 35 demonstrated low risk. A study was classified as having high risk of bias when withdrawals were not adequately explained and exceeded 5% of the enrolled population. In all studies, the interval between index test and reference standard was 12 months or less, and the most commonly reported time interval was up to three months. In every study, all participants received the same reference standard.

One study presented high concern for patient selection applicability (Eskenazi 2001), and the remaining 48 studies demonstrated low concern. We assigned high concern for patient selection applicability if the study utilised two-gate selection for cases and controls, as any sampling deviation from a representative group of the entire clinically relevant population could skew the estimates of diagnostic accuracy in any direction. No studies had concerns about applicability in 'index test' and 'reference standard' domains.

Findings

Findings are presented under two main categories.

- Diagnostic tests for endometriosis (Summary of findings 1).
- Mapping of DIE to specific anatomical sites (Summary of findings 2).

Diagnostic tests for endometriosis

We analysed the diagnostic test accuracy of imaging tests for three types of endometriosis in a total of 29 studies.

- Pelvic endometriosis at all locations at any depth of invasion (13 studies, 1535 participants).
- Ovarian endometriosis (10 studies, 852 participants).
- DIE/posterior DIE (15 studies, 1493 participants).

Findings are outlined in Summary of findings 1 and Appendix 2. Twelve studies performed eight head-to-head direct comparisons between tests. Data were insufficient to permit meta-analyses of paired tests, hence, we have reported available comparative results narratively and have illustrated them in forest plots and ROC plots.

Pelvic endometriosis

Pelvic endometriosis using ultrasonography

Five articles, which included a total of 1222 participants, were published between 2001 and 2014 and explored the accuracy of TVUS in diagnosing pelvic endometriosis. These studies were conducted in Europe ($n = 4$) and in the Middle East ($n = 1$). The mean sensitivity and specificity of all included studies were 0.65 (95% confidence interval (CI) 0.27 to 1.00) and 0.95 (95% CI 0.89 to 1.00). Four studies evaluated conventional TVUS, and one study addressed the tenderness-guided method (tg-TVUS). Forest plots (Figure 5) and the ROC plot (Figure 6) demonstrated a high degree of heterogeneity between papers, which was greater for estimates of sensitivity than of specificity. One of the studies (710 participants) (Ghezzi 2005) utilised the 'kissing sign' as a sole single marker of endometriosis, in contrast to the other four studies, which surveyed pelvic anatomy in general. This paper reported markedly low sensitivity at 0.09 (95% CI 0.06 to 0.12), which influenced the sensitivity estimate for the group, as the range of sensitivities for the other four studies (512 participants) was between 0.56 and 0.96, whereas specificities ranged between 0.80 and 0.99. The mean sensitivity and specificity of these four studies were 0.79 (95% CI 0.36 to 1.00) and 0.91 (95% CI 0.74 to 1.00). Even when data from a large outlying study were excluded, sensitivity and specificity estimates were heterogeneous and confidence intervals wide, and estimates did not meet the criteria for a replacement or a triage test but approached the criteria for a SpPin triage test. No other ultrasound techniques were evaluated as a diagnostic test for pelvic endometriosis.

Figure 5. Forest plot of TVUS for detection of pelvic endometriosis. Plot shows study-specific estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are ordered according to the year of publication. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional TVUS are presented as 'modified method'.

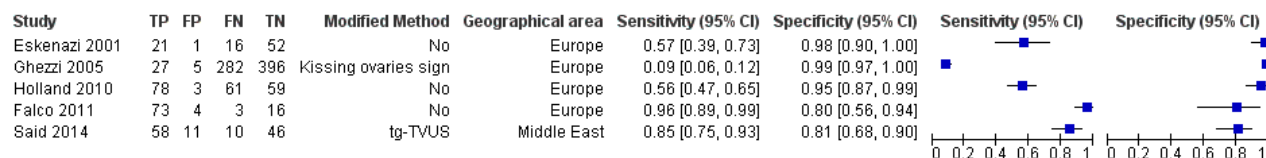
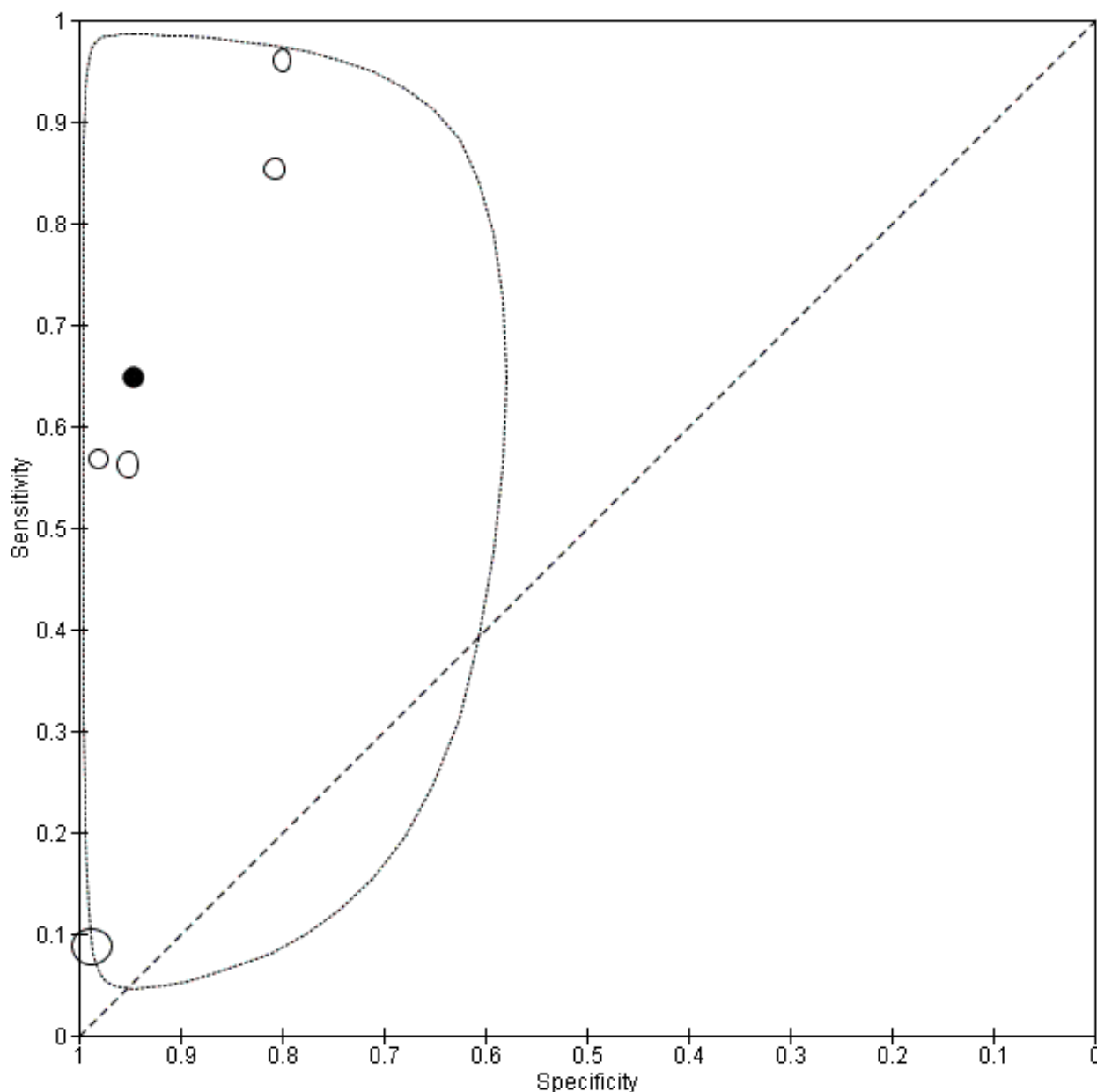


Figure 6. Summary ROC plot of TVUS for detection of pelvic endometriosis. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).



Pelvic endometriosis using MRI

Seven studies, including 10 data sets with a total of 303 participants, assessed the value of MRI in detecting pelvic endometriosis. Eligible MRI evaluations were published between 1993 and 2011, and most ($n = 4$) were published in the early 1990s. Studies were conducted in Asia ($n = 3$), North America ($n = 2$) and Europe ($n = 2$). Five different MRI methods were assessed: (1) T1/T2-w MRI (three studies, 97 participants); (2) fat-suppressed MRI (one study, 31 participants); (3) T1/T2-w MRI with fat-suppression (two

studies, 105 participants); (4) T1/T2-w MRI with fat-suppression/Gd (two studies, 77 participants); and (5) 3.0T MRI (two studies, 86 participants). Three studies compared more than one MRI method in the same cohort of women (Sugimura 1993; Ha 1994; Ascher 1995). The mean sensitivity and specificity of all included studies were 0.79 (95% CI 0.70 to 0.88) and 0.72 (95% CI 0.51 to 0.90), which did not meet the criteria for a replacement or a triage test. Forest plots (Figure 7) and the ROC plot (Figure 8) showed a high degree of heterogeneity for estimates of both sensitivity and specificity.

Figure 7. Forest plot of MRI for detection of pelvic endometriosis. Plot shows study-specific estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are ordered according to year of publication. Tests on the same population (different MRI methods) are presented separately as MRI* and MRI. FN: false negative; FP: false positive; TN: true negative; TP: true positive.**

MRI pelvic

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Sugimura 1993	19	3	7	6	Conventional [T1-/T2-w]	Asia	0.73 [0.52, 0.88]	0.67 [0.30, 0.93]		
Ha 1994	15	0	14	2	Conventional [T1-/T2-w]	Asia	0.52 [0.33, 0.71]	1.00 [0.16, 1.00]		
Ascher 1995	16	4	5	6	Conventional [T1-/T2-w]	North America	0.76 [0.53, 0.92]	0.60 [0.26, 0.88]		
Okada 1995	57	3	8	6	Fat-suppressed	Asia	0.88 [0.77, 0.95]	0.67 [0.30, 0.93]		
Stratton 2003	28	1	14	3	T1/T2-w + fat-suppressed/Gd	North America	0.67 [0.50, 0.80]	0.75 [0.19, 0.99]		
Manganaro 2012a	31	0	1	14	3.0T	Europe	0.97 [0.84, 1.00]	1.00 [0.77, 1.00]		
Thomeer 2014	30	0	7	3	3.0T	Europe	0.81 [0.65, 0.92]	1.00 [0.29, 1.00]		

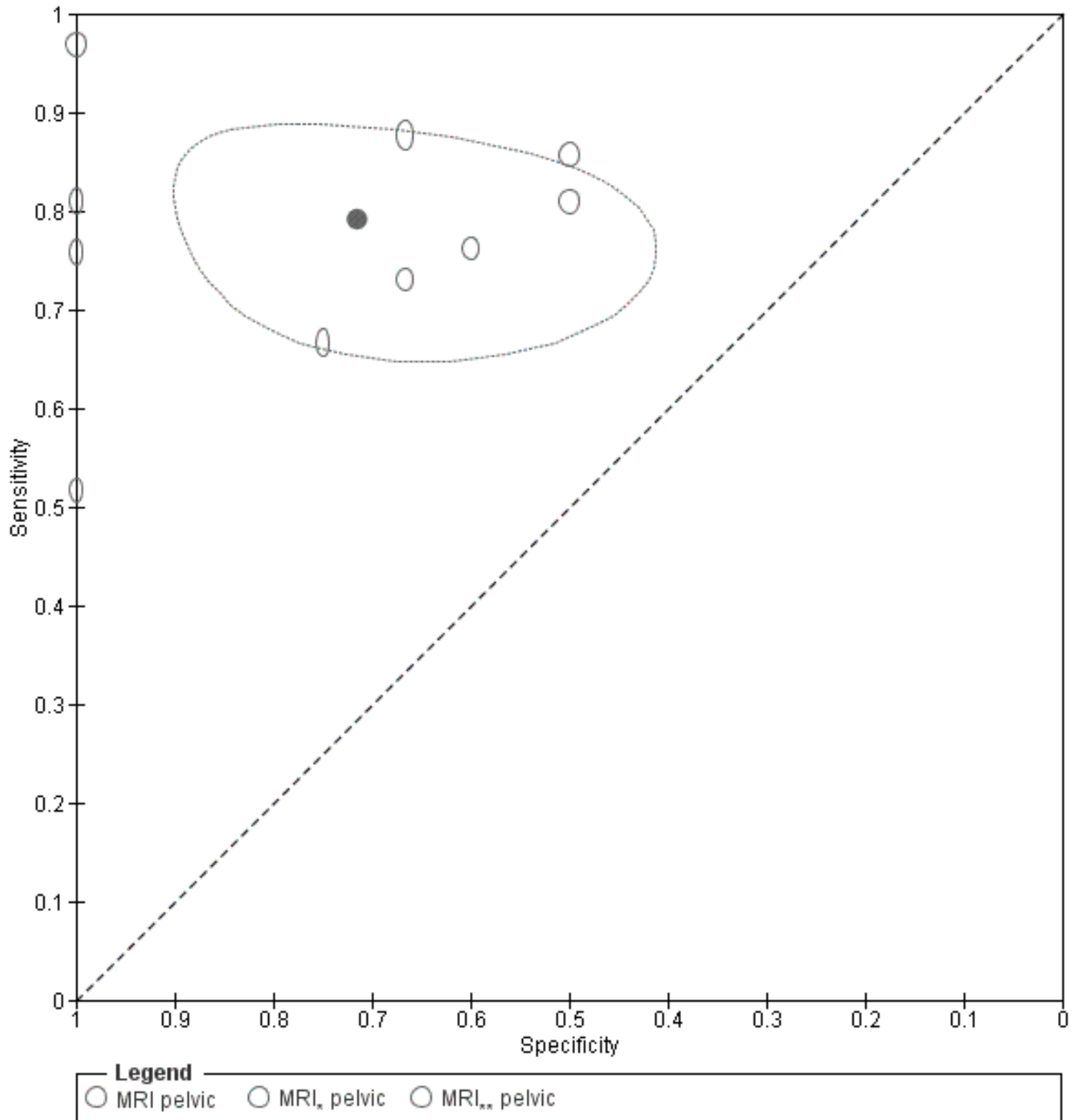
MRI* pelvic

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Ha 1994	22	0	7	2	Fat-suppressed	Asia	0.76 [0.56, 0.90]	1.00 [0.16, 1.00]		
Ascher 1995	18	5	3	5	T1/T2-w + fat-suppressed	North America	0.86 [0.64, 0.97]	0.50 [0.19, 0.81]		

MRI pelvic**

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Ascher 1995	17	5	4	5	T1/T2-w + fat-suppressed/Gd	North America	0.81 [0.58, 0.95]	0.50 [0.19, 0.81]		

Figure 8. Summary ROC plot of MRI for detection of pelvic endometriosis. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. Tests on the same population (different MRI methods) are presented separately as MRI* and MRI. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).**



Pelvic endometriosis using other imaging modalities

Authors of one paper determined the accuracy of ^{18}F FDG PET-CT in detecting pelvic endometriosis (10 participants, published in 2011, conducted in Europe), showing sensitivity of 0.00 (95% CI 0.00 to 0.34) and specificity of 1.00 (95% CI 0.03 to 1.00). Similarly, different groups in another small descriptive study showed negative findings for the same test; this study did not meet the inclusion criteria (Setubal 2011). No other imaging techniques described in the included studies evaluated pelvic endometriosis.

Indirect comparisons of imaging tests for pelvic endometriosis

With regards to TVUS modalities, no specific technique, year of publication or geographical location resulted in a better performing method. The two most recent small studies evaluated 3.0T MRI; each showed high sensitivity and specificity for diagnosing pelvic endometriosis (sensitivity 0.97, 95% CI 0.84 to 1.00; specificity 1.00,

95% CI 0.77 to 1.00 - Manganaro 2012a; sensitivity 0.81, 95% CI 0.65 to 0.92; specificity 1.00, 95% CI 0.29 to 1.00 - Thomeer 2014). The latter study displayed wide confidence intervals, suggesting that caution should be used in interpreting these findings. Different MRI methods were not formally compared because the small number of studies and their small size precluded meaningful results.

Mean estimates of TVUS after exclusion of the outlier study showed comparable sensitivity but higher specificity than were seen with MRI.

Direct comparisons of imaging tests for pelvic endometriosis

Three studies made a direct head-to-head comparison of two or three MRI methods, but all were small and inconclusive and reported wide and overlapping confidence intervals (Ha 1994; Ascher 1995; Sugimura 1993) (see Appendix 2; Figure 9; Figure 10; Figure 11). No studies have compared MRI and TVUS.

Figure 9. Forest plot demonstrating the direct comparison between MRI methods for pelvic endometriosis. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

MRI pelvic

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Ascher 1995	16	4	5	6	Conventional [T1-/T2-w]	North America	0.76 [0.53, 0.92]	0.60 [0.26, 0.88]		
Ha 1994	15	0	14	2	Conventional [T1-/T2-w]	Asia	0.52 [0.33, 0.71]	1.00 [0.16, 1.00]		

MRI, pelvic

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Ascher 1995	18	5	3	5	T1/T2-w + fat-suppressed	North America	0.86 [0.64, 0.97]	0.50 [0.19, 0.81]		
Ha 1994	22	0	7	2	Fat-suppressed	Asia	0.76 [0.56, 0.90]	1.00 [0.16, 1.00]		

Figure 10. Forest plot demonstrating the direct comparison between MRI methods for pelvic endometriosis. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

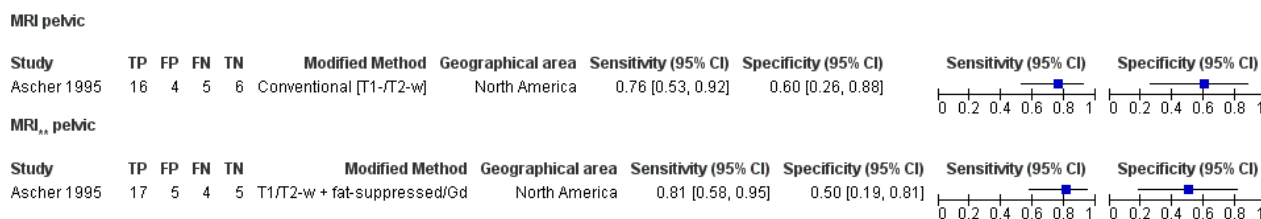
MRI pelvic

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Sugimura 1993	19	3	7	6	Conventional [T1-/T2-w]	Asia	0.73 [0.52, 0.88]	0.67 [0.30, 0.93]		

MRI pelvic¹

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Sugimura 1993	20	2	6	7	Fat-suppressed	Asia	0.77 [0.56, 0.91]	0.78 [0.40, 0.97]		

Figure 11. Forest plot demonstrating the direct comparison between MRI methods for pelvic endometriosis. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive.



Ovarian endometriosis

Ovarian endometriosis using ultrasonography

Eight studies with a total of 765 participants explored the diagnostic accuracy of TVUS for ovarian endometriosis. These included studies were published between 1996 and 2015. Studies were conducted in Europe (n = 6), Australia (n = 1) and South America (n = 1).

Mean sensitivity and specificity estimates for all included studies were 0.93 (95% CI 0.87 and 0.99) and 0.96 (95% CI 0.92 and 0.99), respectively, meeting the criteria for a SpPin triage test and approaching the criteria for a replacement tet and a SnNout triage test. Estimates for both sensitivity and specificity showed less heterogeneity than were seen in other types of endometriosis (Figure 12; Figure 13).

Figure 12. Forest plot of US methods (TVUS, TRUS) for detection of ovarian endometriosis. Plot shows study-specific estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are presented for TVUS and TRUS and are ordered according to year of publication. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional TVUS are presented as 'modified method'.

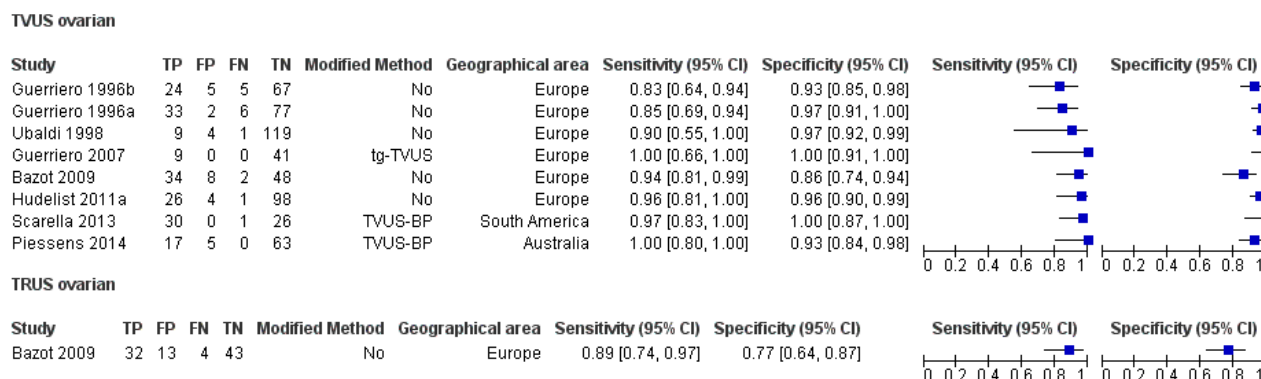
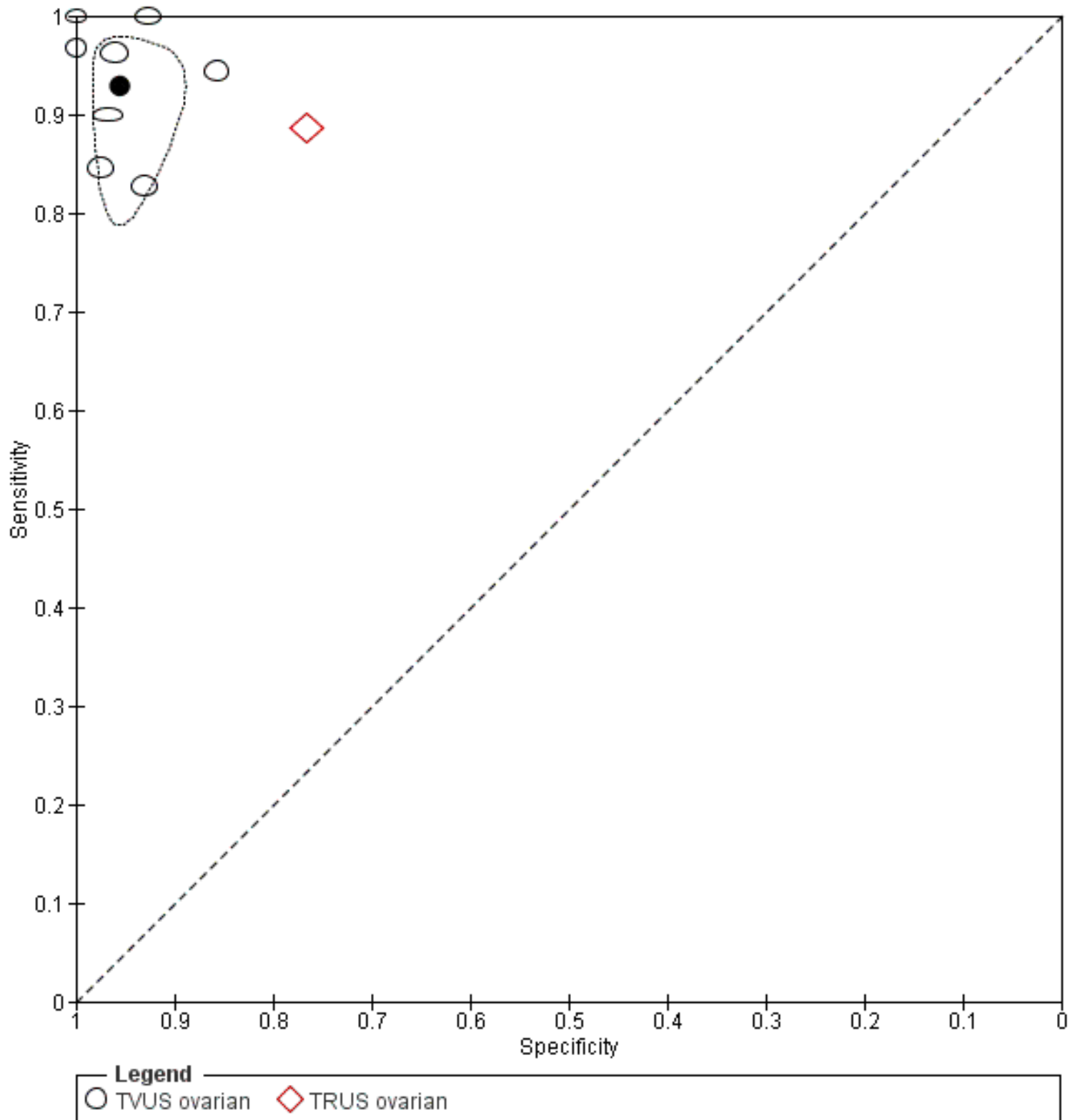


Figure 13. Summary ROC plot of US methods (TVUS, TRUS) for detection of ovarian endometriosis. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study

sample size. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line) (for TVUS).



Ovarian endometriosis using MRI

Three studies with a total of 179 participants were published in 2009 and 2011 and assessed the diagnostic accuracy of MRI for ovarian endometriosis. All studies were conducted in Europe. One study (92 participants) used T1/T2-w MRI with fat-suppression/Gd, and two studies (87 participants) utilised 3.0T MRI. Meta-analysis

of these three studies revealed summary sensitivity and specificity of 0.95 (95% CI 0.90 to 1.00) and 0.91 (95% CI 0.86 to 0.97), meeting the criteria for a replacement test and a SnNout triage test, and approaching the criteria for a SpPin triage test (Figure 14). However, the few identified studies provided insufficient evidence to allow meaningful conclusions on the diagnostic role of MRI for endometrioma.

Figure 14. Forest plot of MRI for detection of ovarian endometriosis. Plot shows study-specific estimates of sensitivity and specificity (squares) with 95% CI (black line). Studies are ordered by year of publication. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	33	7	3	49	T1/T2-w + fat-suppressed/Gd	Europe	0.92 [0.78, 0.98]	0.88 [0.76, 0.95]		
Hottat 2009	20	1	1	19	3.0T	Europe	0.95 [0.76, 1.00]	0.95 [0.75, 1.00]		
Manganaro 2012a	19	1	0	26	3.0T	Europe	1.00 [0.82, 1.00]	0.96 [0.81, 1.00]		

Indirect comparisons of imaging tests for ovarian endometriosis

For TVUS, articles published after 2006 ($n = 5$) demonstrated higher sensitivity for diagnosing endometrioma. The most accurate ultrasound methods appeared to be tenderness-guided TVUS (one study in 50 women), which showed sensitivity of 1.00 (95% CI 0.66 to 1.00) and specificity of 1.00 (95% CI 0.91 to 1.00) (Guerrero 2007), and TVUS-BP (two studies in 142 women), which demonstrated sensitivity of 0.97 (95% CI 0.83 to 1.00) and 1.00 (95% CI 0.81 to 1.00) and specificity of 1.00 (95% CI 0.87 to 1.00) and 0.93 (95% CI 0.84 to 0.98) (Scarella 2013; Piessens 2014). Data were insufficient to permit formal comparisons of TVUS methods.

Higher estimates were reported for 3.0T MRI with sensitivities of 0.95 and 1.00 (95% CI 0.76 to 1.00 and 0.82 to 1.00) and specificities of 0.95 and 0.96 (95% CI 0.75 to 1.00 and 0.81 to 1.00) than for T1/T2-w MRI with fat-suppression/Gd, which showed sensitivity of 0.92

(95% CI 0.78 to 0.98) and specificity of 0.88 (95% CI 0.76 and 0.95), although confidence intervals overlapped.

When pooled estimates were considered, TVUS showed lower sensitivity but higher specificity compared with MRI.

Direct comparisons of imaging tests for ovarian endometriosis

One study (92 participants, published in 2009, conducted in Europe) evaluated TRUS and demonstrated sensitivity of 0.89 (95% CI 0.74 to 0.97) and specificity of 0.77 (95% CI 0.64 to 0.84) for diagnosis of ovarian endometriosis (Figure 12). This study directly compared TRUS, TVUS and MRI (Bazot 2009) and found that TRUS had lower diagnostic estimates than TVUS (sensitivity 0.94, 95% CI 0.81 to 0.99; specificity 0.86, 95% CI 0.74 to 0.94) and MRI (sensitivity 0.92, 95% CI 0.78 to 0.98; specificity 0.88, 95% CI 0.76 to 0.95). TVUS and MRI provided comparable estimates for diagnosing ovarian endometriosis (Appendix 3: Figure 15; Figure 16; Figure 17).

Figure 15. Forest plot demonstrating the direct comparison between TVUS and TRUS for ovarian endometriosis. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS ovarian										
Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	34	8	2	48	No	Europe	0.94 [0.81, 0.99]	0.86 [0.74, 0.94]		
TRUS ovarian										
Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	32	13	4	43	No	Europe	0.89 [0.74, 0.97]	0.77 [0.64, 0.87]		

Figure 16. Forest plot demonstrating the direct comparison between TRUS and MRI for ovarian endometriosis. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

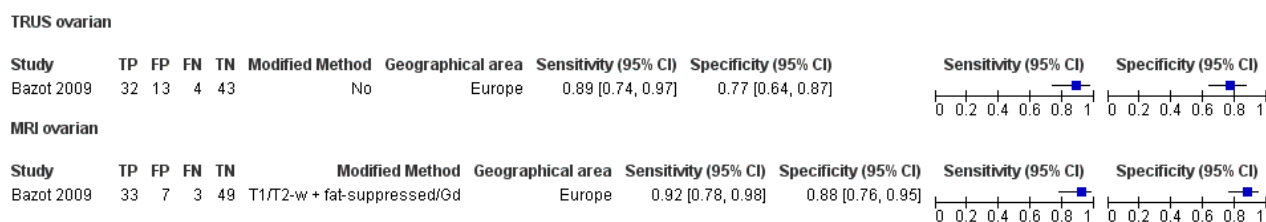
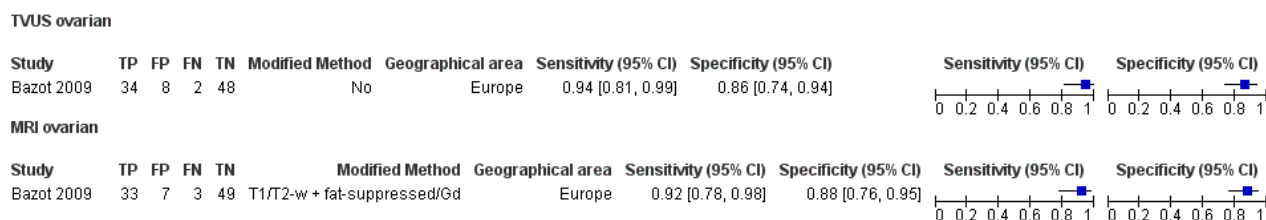


Figure 17. Forest plot demonstrating the direct comparison between TVUS and MRI for ovarian endometriosis. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.



We identified no comparative studies of MRI and other imaging tests for ovarian endometriosis, other than the one presented above.

Deep infiltrating endometriosis/posterior DIE

Deep infiltrating endometriosis using ultrasonography

Nine articles included 12 data sets with a total of 934 participants and assessed the accuracy of TVUS in detecting DIE ($n = 3$) and posterior DIE ($n = 7$). All included studies were published after 2002, and most ($n = 7$) were published after 2009. These studies were

conducted in Europe ($n = 7$), South America ($n = 1$) and Australia ($n = 1$). TVUS techniques included (1) TVUS (seven studies, eight data sets, 721 participants); (2) 3D-TVUS (two studies, 226 participants); and (3) SVG (two studies, 235 participants). Mean sensitivity and specificity estimates for all included studies were 0.79 (95% CI 0.69 to 0.89) and 0.94 (95% CI 0.88 to 1.00), which approached the criteria for a SpPin triage test. Forest plots (Figure 18) and the ROC plot (Figure 19) revealed a high degree of heterogeneity for both sensitivity and specificity, with greater heterogeneity for sensitivity.

Figure 18. Forest plot of TVUS for detection of DIE/Posterior DIE. Plot shows study-specific estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are ordered according to year of publication for DIE and Posterior DIE, respectively. Tests on the same population (different TVUS methods) are presented separately as TVUS*. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional TVUS are presented as 'modified method'.

TVUS DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Holland 2010	23	7	15	156	No	Europe	0.61 [0.43, 0.76]	0.96 [0.91, 0.98]		
Grasso 2010	15	2	4	3	3D-TVUS	Europe	0.79 [0.54, 0.94]	0.60 [0.15, 0.95]		
Scarella 2013	33	0	2	22	TVUS-BP	South America	0.94 [0.81, 0.99]	1.00 [0.85, 1.00]		

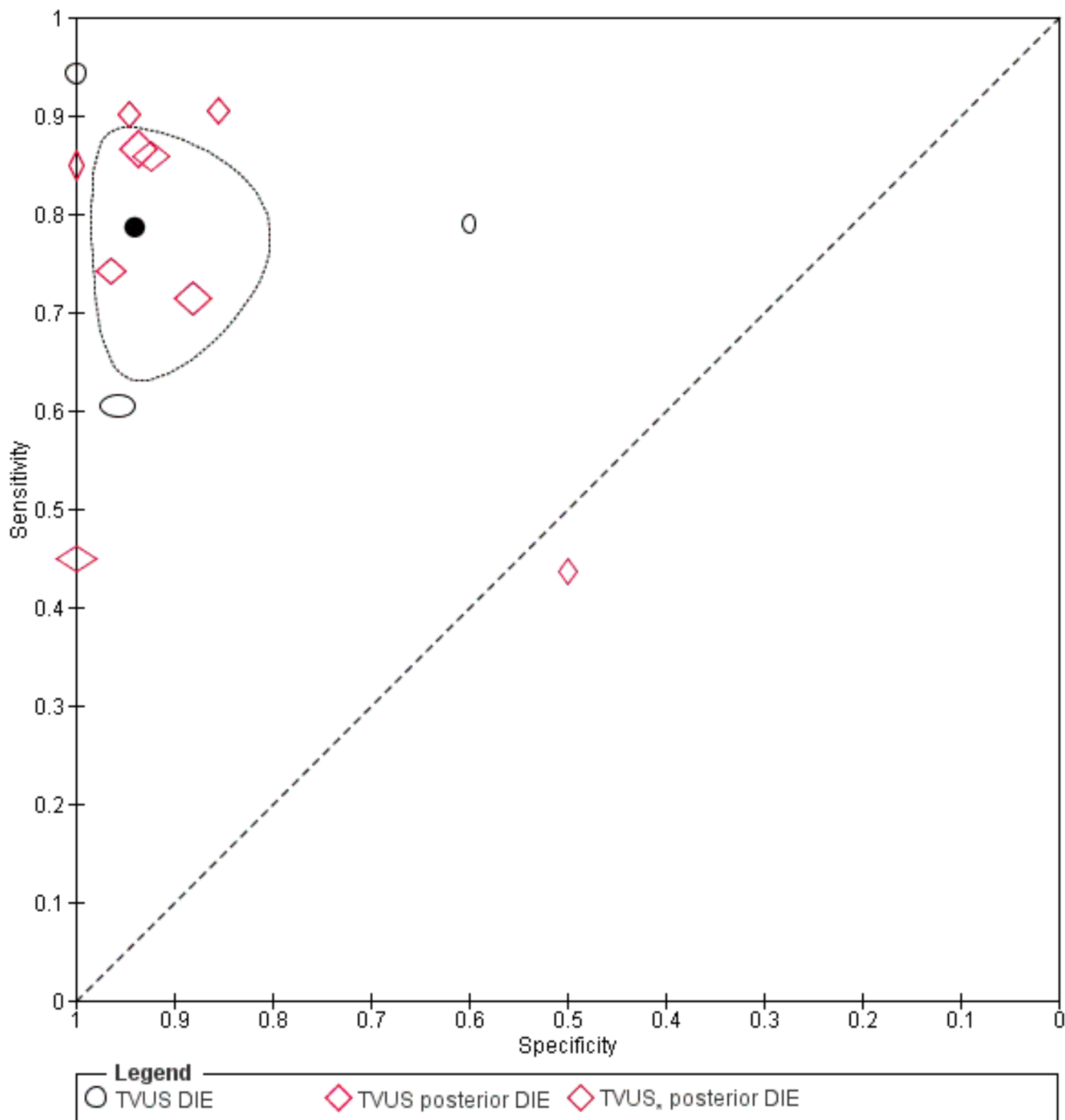
TVUS posterior DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Dessole 2003	14	7	18	7	No	Europe	0.44 [0.26, 0.62]	0.50 [0.23, 0.77]		
Guerriero 2007	28	1	3	18	tg-TVUS	Europe	0.90 [0.74, 0.98]	0.95 [0.74, 1.00]		
Holland 2010	14	0	17	170	No	Europe	0.45 [0.27, 0.64]	1.00 [0.98, 1.00]		
Savelli 2011	57	0	10	2	No	Europe	0.85 [0.74, 0.93]	1.00 [0.16, 1.00]		
Falco 2011	29	2	10	55	No	Europe	0.74 [0.58, 0.87]	0.96 [0.88, 1.00]		
Reid 2014	49	10	8	122	SVG	Australia	0.86 [0.74, 0.94]	0.92 [0.87, 0.96]		
Guerriero 2014	65	13	26	98	tg-TVUS	Europe	0.71 [0.61, 0.80]	0.88 [0.81, 0.94]		

TVUS, posterior DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Dessole 2003	29	2	3	12	SVG	Europe	0.91 [0.75, 0.98]	0.86 [0.57, 0.98]		
Guerriero 2014	79	7	12	104	3D-TVUS	Europe	0.87 [0.78, 0.93]	0.94 [0.87, 0.97]		

Figure 19. Summary ROC plot of TVUS for detection of DIE/Posterior DIE. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. Tests on the same population (different TVUS methods) are presented separately as TVUS_s. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).



Deep infiltrating endometriosis using MRI

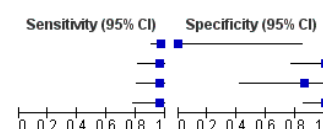
Six studies, including seven data sets with a total of 266 participants, evaluated MRI for the diagnosis of DIE (n = 4) and posterior DIE (n = 2; three data sets). All studies were published after 2004 and were conducted in Europe (n = 5) and Asia (n = 1). MRI methods included (1) MRI jelly (one study, 31 participants); (2) T1/T2-w MRI with fat-suppression/Gd (two studies, 125 participants);

(3) 2D-MRI T2-w (one study, 23 participants); (4) 3D-MRI (one study, 23 participants); and (5) 3.0T MRI (two studies, 87 participants). Mean estimates of sensitivity and specificity for all studies were 0.94 (95% CI 0.90 to 0.97) and 0.77 (95% CI 0.44 to 1.00), which approached the criteria for a replacement test and a SnNout triage test. Forest plots (Figure 20) and the ROC plot (Figure 21) showed greater heterogeneity for estimates of specificity than sensitivity.

Figure 20. Forest plot of MRI for detection of DIE/Posterior DIE. Plot shows study-specific estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are ordered according to year of publication. Tests on the same population (different MRI methods) are presented separately as MRI*. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional MRI are presented as 'modified method'.

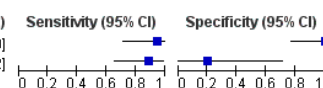
MRI DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	87	2	3	0	T1/T2-w + fat-suppressed/Gd	Europe	0.97 [0.91, 0.99]	0.00 [0.00, 0.84]
Hottat 2009	26	0	1	14	3.0T	Europe	0.96 [0.81, 1.00]	1.00 [0.77, 1.00]
Grasso 2010	25	1	1	6	T1/T2-w + fat-suppressed/Gd	Europe	0.96 [0.80, 1.00]	0.86 [0.42, 1.00]
Manganaro 2012a	22	0	1	23	3.0T	Europe	0.96 [0.78, 1.00]	1.00 [0.85, 1.00]



MRI posterior DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Takeuchi 2005	16	0	1	14	Jelly method [T1/T2-w + fat-suppressed]	Asia	0.94 [0.71, 1.00]	1.00 [0.77, 1.00]
Bazot 2013	16	4	2	1	2D FSE T2-w	Europe	0.89 [0.65, 0.99]	0.20 [0.01, 0.72]



MRI, posterior DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2013	18	4	0	1	3D-MRI	Europe	1.00 [0.81, 1.00]	0.20 [0.01, 0.72]

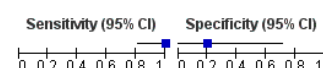
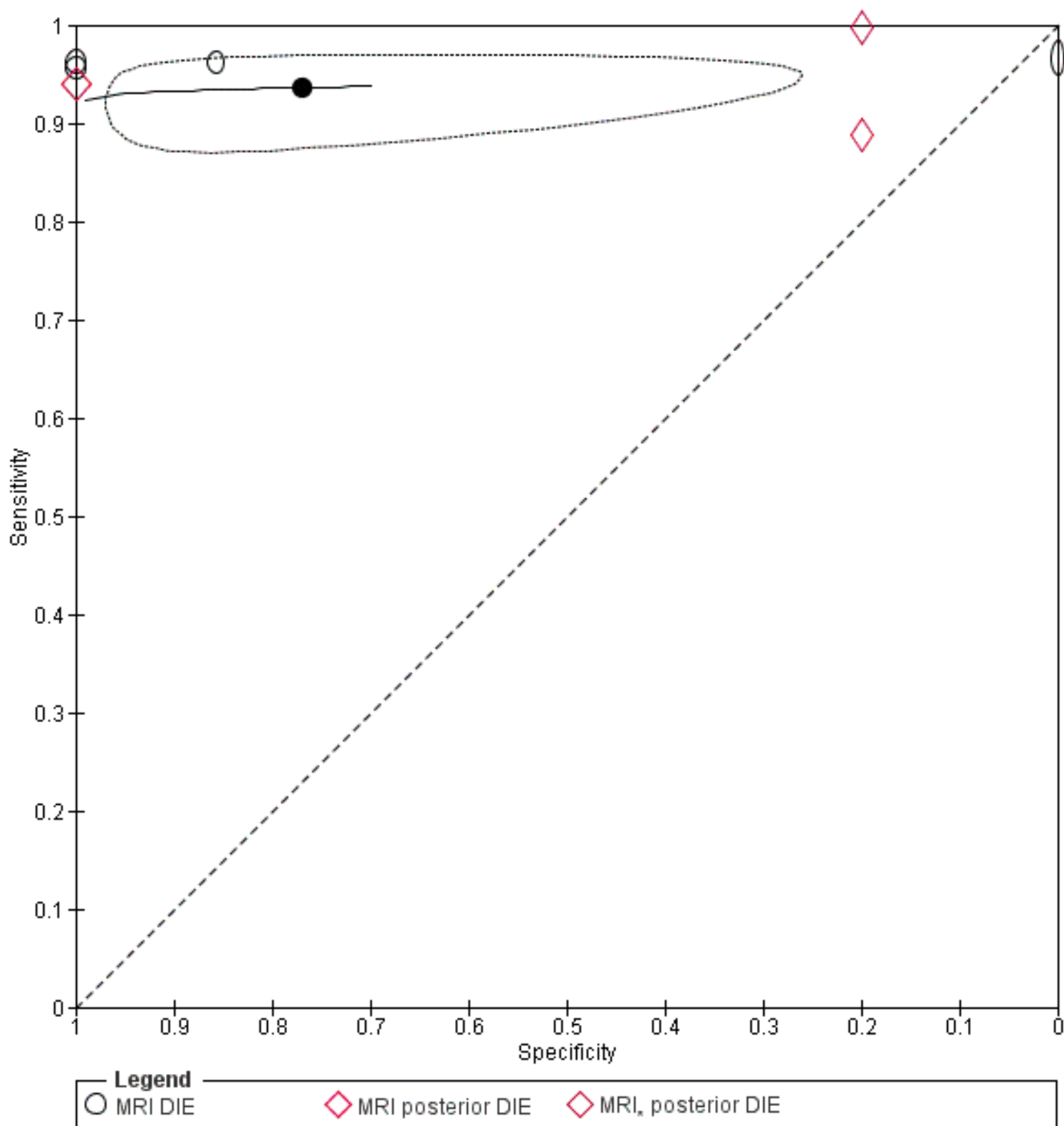


Figure 21. Summary ROC plot of MRI for detection of DIE/Posterior DIE. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. Tests on the same

population (different MRI methods) are presented separately as MRI*. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).



Deep infiltrating endometriosis using other imaging modalities

One study determined the accuracy of double-contrast barium enema (DCBE) in detecting DIE (69 participants, published in 2011, conducted in Europe), showing sensitivity of 0.36 (95% CI 0.24 to

0.48) and specificity of 1.00 (95% CI 0.16 to 1.00). This test was inferior to TVUS when directly compared in the same individuals (Appendix 4: Figure 22). The included studies evaluated no other imaging techniques for DIE/posterior DIE.

Figure 22. Forest plot demonstrating the direct comparison between TVUS and DCBE for DIE/Posterior DIE. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS posterior DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Savelli 2011	57	0	10	2	No	Europe	0.85 [0.74, 0.93]	1.00 [0.16, 1.00]		

DCBE DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Savelli 2011	24	0	43	2	No	Europe	0.36 [0.24, 0.48]	1.00 [0.16, 1.00]		

Indirect comparisons of imaging tests for deep infiltrating endometriosis

TVUS-BP (one study, 57 participants) (Scarella 2013) showed the highest diagnostic accuracy of all TVUS methods with sensitivity of 0.94 (95% CI 0.81 to 0.99) and specificity of 1.00 (95% CI 0.85 to 1.00). Tenderness-guided TVUS (one study, 50 participants) (Guerriero 2007) had relatively high sensitivity of 0.90 (95% CI 0.74 to 0.98) and high specificity of 0.95 (95% CI 0.74 to 1.00), but a subsequent study by the same group using the same methods in a separate cohort (172 participants) (Guerriero 2014) did not reach a similar level of diagnostic accuracy with sensitivity of 0.71 (95% CI 0.61 to 0.80) and specificity of 0.88 (95% CI 0.81 to 0.94). Data were insufficient for a formal comparison of different methods of TVUS. Researchers evaluated no other ultrasound techniques as a diagnostic test for DIE/posterior DIE.

3.0T MRI (Hottat 2009; Manganaro 2012a) showed the highest diagnostic accuracy with sensitivity of 0.96 (95% CI 0.78 to 1.00 and 0.81 to 1.00) and specificity of 1.00 (95% CI 0.77 to 1.00 and 0.85 to 1.00), and the MRI jelly method (Takeuchi 2005) with sensitivity of 0.94 (95% CI 0.71 to 1.00) and specificity of 1.00 (95% CI 0.77 to 1.00). Data were insufficient for formal comparative analyses between MRI methods for DIE/posterior DIE.

Similarly to ovarian endometriosis, pooled estimates of TVUS demonstrated lower sensitivity but higher specificity compared with MRI.

Direct comparisons of imaging tests for deep infiltrating endometriosis

- Direct comparison between tenderness-guided TVUS and 3D-TVUS (one study, 202 participants) (Guerriero 2014) revealed that conventional TVUS is less accurate (sensitivity 0.71, 95% CI 0.61 to 0.80; specificity 0.88, 95% CI 0.81 to 0.94) than 3D-TVUS (sensitivity 0.87, 95% CI 0.78 to 0.93; specificity 0.94, 95% CI 0.87 to 0.97) (Appendix 4: Figure 23).
- TVUS had lower estimates of sensitivity 0.44 (95% CI 0.26 to 0.62) and specificity 0.50 (95% CI 0.23 to 0.77) compared with SVG (sensitivity 0.91, 95% CI 0.75 to 0.98; specificity 0.86, 95% CI 0.57 to 0.98) in another study of 46 women (Dessole 2003) (Appendix 4: Figure 23).
- One paired evaluation (23 participants) (Bazot 2013) demonstrated that 3D-MRI had higher sensitivity (1.0, 95% CI 0.81 to 1.00) than 2D-MRI (0.89, 95% CI 0.65 to 0.99), but both tests had identically low specificity of 0.2 (95% CI 0.01 to 0.72) (Appendix 4: Figure 24).
- MRI (sensitivity 0.96, 95% CI 0.80 to 1.00; specificity 0.86, 95% CI 0.42 to 1.00) appeared to be superior to 3D-TVUS (sensitivity 0.79, 95% CI 0.54 to 0.94; specificity 0.60, 95% CI 0.15 to 0.95) in one small study that had unequal numbers of participants (MRI, n = 33; 3D-TVUS, n = 25) from the same cohort (Grasso 2010) (Appendix 4: Figure 25).

Figure 23. Forest plot demonstrating the direct comparison between TVUS methods for DIE/Posterior DIE. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS posterior DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Dessole 2003	14	7	18	7	No	Europe	0.44 [0.26, 0.62]	0.50 [0.23, 0.77]		
Guerriero 2014	65	13	26	98	tg-TVUS	Europe	0.71 [0.61, 0.80]	0.88 [0.81, 0.94]		

TVUS, posterior DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Dessole 2003	29	2	3	12	SVG	Europe	0.91 [0.75, 0.98]	0.86 [0.57, 0.98]		
Guerriero 2014	79	7	12	104	3D-TVUS	Europe	0.87 [0.78, 0.93]	0.94 [0.87, 0.97]		

Figure 24. Forest plot demonstrating the direct comparison between MRI methods for DIE/Posterior DIE. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

MRI posterior DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2013	16	4	2	1	2D FSE T2-w	Europe	0.89 [0.65, 0.99]	0.20 [0.01, 0.72]		

MRI, posterior DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2013	18	4	0	1	3D-MRI	Europe	1.00 [0.81, 1.00]	0.20 [0.01, 0.72]		

Figure 25. Forest plot demonstrating the direct comparison between 3D-TVUS and MRI for DIE/Posterior DIE. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Grasso 2010	15	2	4	3	3D-TVUS	Europe	0.79 [0.54, 0.94]	0.60 [0.15, 0.95]		

MRI DIE

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Grasso 2010	25	1	1	6	T1/T2-w + fat-suppressed/Gd	Europe	0.96 [0.80, 1.00]	0.86 [0.42, 1.00]		

Mapping of DIE to specific anatomical sites

A total of 33 studies evaluated the ability of imaging tests to accurately map endometriotic lesions to specific anatomical sites within the pelvic cavity (see [Target conditions](#)). Most papers described more than one anatomical site and/or assessed more than one imaging test. Ninety-four per cent (31/33) were published between 2007 and 2015 ([Summary of findings 2](#) Appendix 6). Twenty-seven studies reported a total of 25 direct imaging modality

comparisons in mapping endometriotic lesions. Insufficient data and considerable concerns about the risk of bias undermined the validity and reliability of results obtained from these comparisons. Study-level comparative data are presented in a descriptive form for each anatomical site.

USL endometriosis

Eleven studies (14 data sets) assessed the diagnostic accuracy of TVUS, TRUS and MRI for detecting USL endometriosis in Europe (n

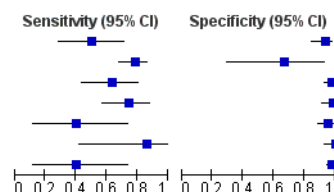
= 8), Australia (n = 2) and South America (n = 1). For TVUS (seven studies, 751 participants), mean sensitivity and specificity were 0.64 (95% CI 0.50 to 0.79) and 0.97 (95% CI 0.93 to 1.00). For MRI (four studies, five data sets, 199 participants), mean sensitivity and specificity were 0.86 (95% CI 0.80 to 0.92) and 0.84 (95% CI 0.68 to 1.00). In the two studies that evaluated TRUS in 232 participants,

summary sensitivity was 0.52 (95% CI 0.29 and 0.74) and summary specificity was 0.94 (95% CI 0.86 to 1.00). For TVUS, estimates of sensitivity were more heterogeneous than those for specificity (Figure 26; Figure 27; Figure 28), whereas for MRI, specificity was more heterogeneous than sensitivity. For TRUS, both sensitivity and specificity were highly variable.

Figure 26. Forest plot of all imaging tests for diagnosis of USL involvement by endometriosis. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are ordered according to year of publication for each test. Tests on the same population (different MRI methods) are presented separately as MRI*. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

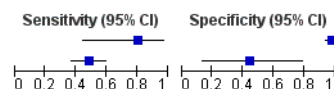
TVUS USL

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 2008	12	4	12	60	tg-TVUS	Europe	0.50 [0.29, 0.71]	0.94 [0.85, 0.98]
Bazot 2009	65	3	18	6	No	Europe	0.78 [0.68, 0.87]	0.67 [0.30, 0.93]
Hudelist 2011a	19	2	11	97	No	Europe	0.63 [0.44, 0.80]	0.98 [0.93, 1.00]
Falco 2011	26	1	9	60	No	Europe	0.74 [0.57, 0.88]	0.98 [0.91, 1.00]
Reid 2013a	4	4	6	86	No	Australia	0.40 [0.12, 0.74]	0.96 [0.89, 0.99]
Scarella 2013	6	0	1	50	TVUS-BP	South America	0.86 [0.42, 1.00]	1.00 [0.93, 1.00]
Reid 2014	4	4	6	175	SVG	Australia	0.40 [0.12, 0.74]	0.98 [0.94, 0.99]



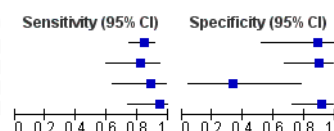
TRUS USL

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Fedele 1998	8	3	2	127	No	Europe	0.80 [0.44, 0.97]	0.98 [0.93, 1.00]
Bazot 2009	40	5	43	4	No	Europe	0.48 [0.37, 0.59]	0.44 [0.14, 0.79]



MRI USL

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	70	1	13	8	T1/T2-w + fat-suppressed/Gd	Europe	0.84 [0.75, 0.91]	0.89 [0.52, 1.00]
Hottat 2009	18	2	4	17	3.0T	Europe	0.82 [0.60, 0.95]	0.89 [0.67, 0.99]
Bazot 2013	15	4	2	2	2D FSE T2-w	Europe	0.88 [0.64, 0.99]	0.33 [0.04, 0.78]
Manganaro 2013	18	2	1	21	3.0T	Europe	0.95 [0.74, 1.00]	0.91 [0.72, 0.99]



MRI_s USL

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2013	15	4	2	2	3D-MRI	Europe	0.88 [0.64, 0.99]	0.33 [0.04, 0.78]

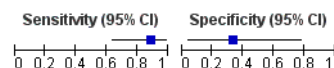


Figure 27. Summary ROC plot of TVUS for detection of USL involvement by endometriosis. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).

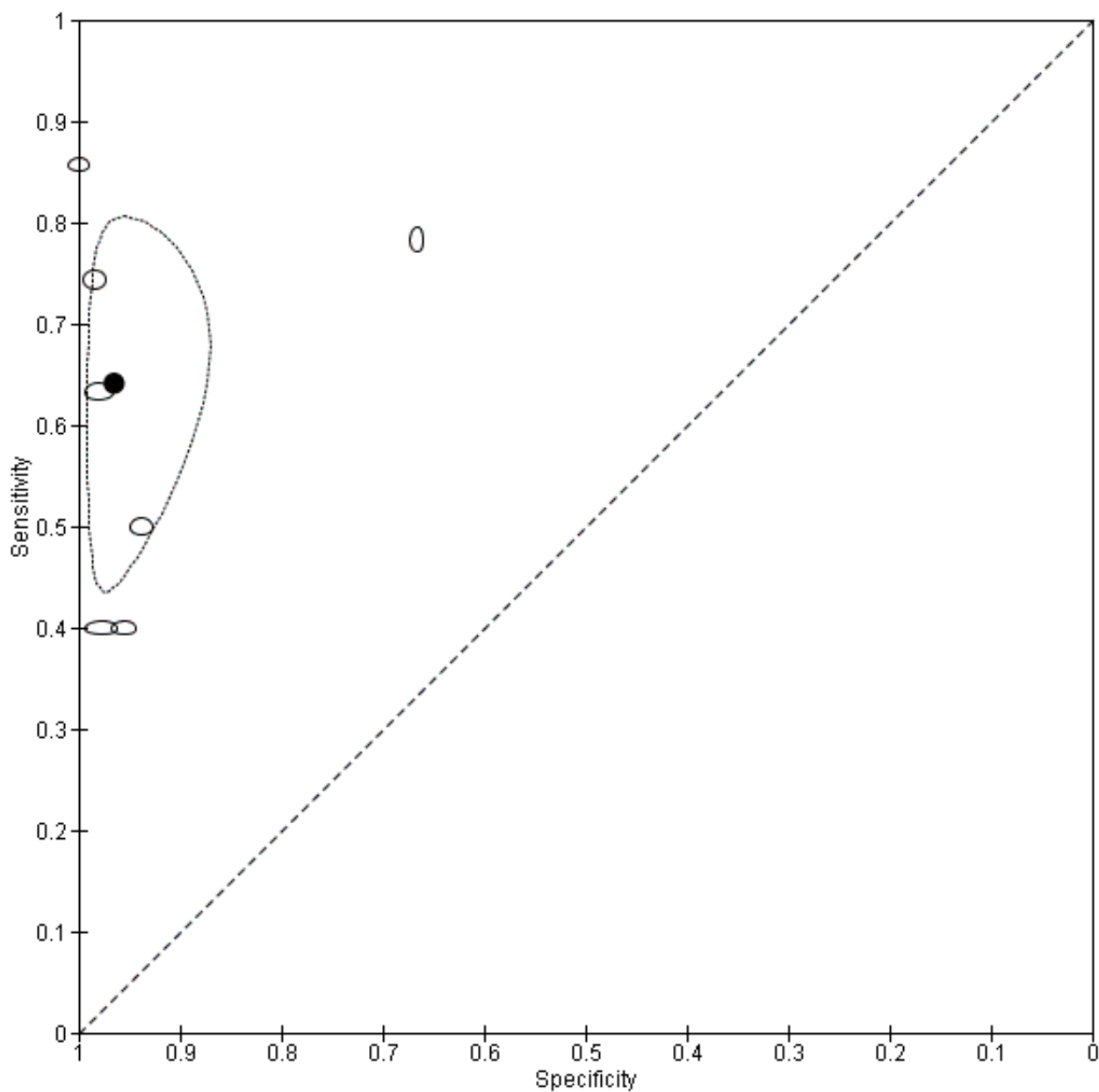
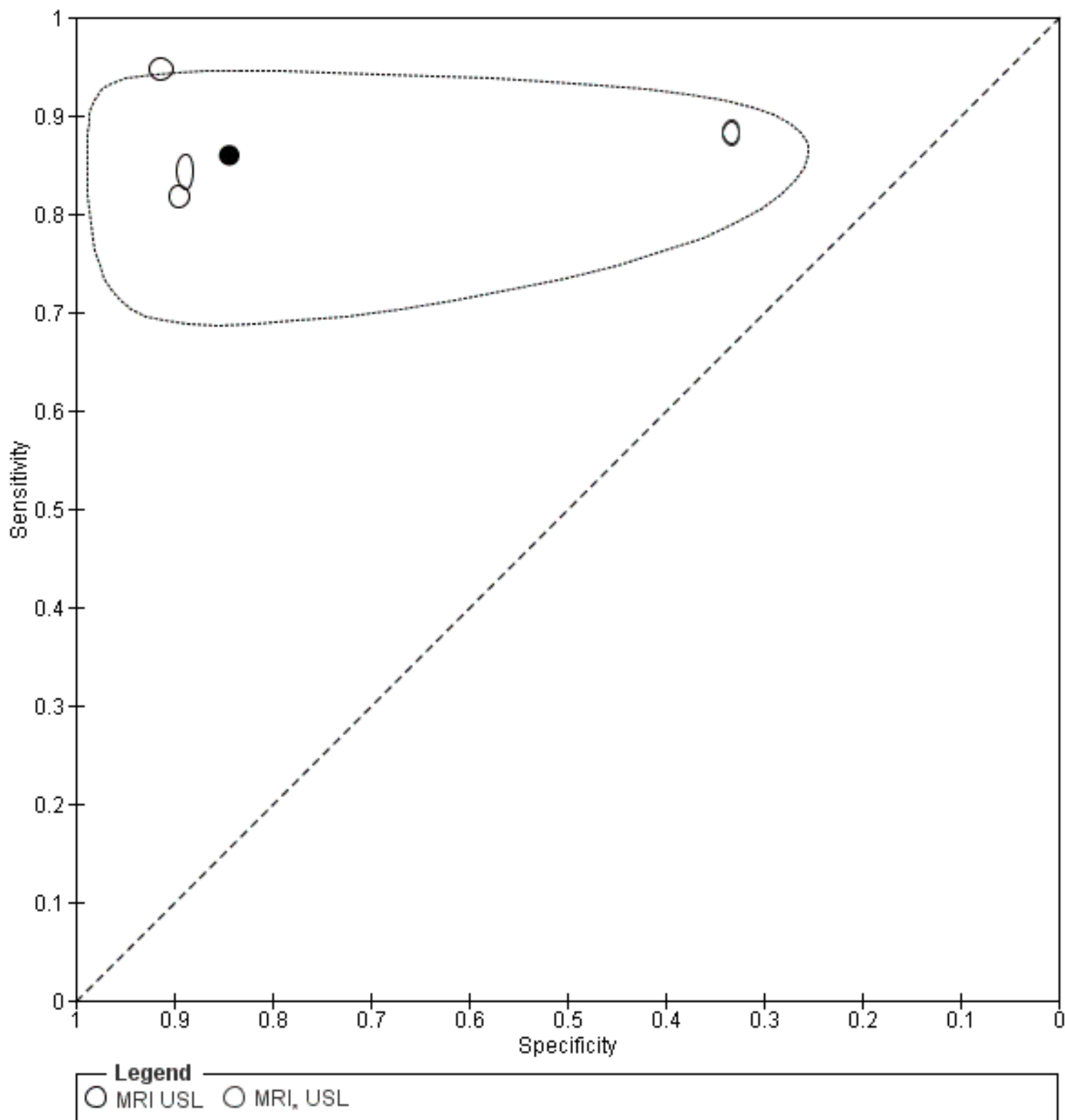


Figure 28. Summary ROC plot of MRI for detection of USL involvement by endometriosis. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. Tests on the same population (different MRI methods) are presented separately as MRI*. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).



Indirect comparisons of imaging tests for USL endometriosis

TVUS-BP (one study, 57 participants) demonstrated the highest sensitivity (0.86, 95% CI 0.42 to 1.00) and specificity (1.00, 95% CI 0.93 to 1.00) of the TVUS methods (Scarella 2013). In the MRI group, the 3.0T method appeared to be highly sensitive (0.95, 95% CI 0.74 to 1.00) and specific (0.91, 95% CI 0.72 to 0.99) in one study that included 42 participants (Manganaro 2013), but it yielded lower diagnostic estimates (sensitivity 0.82, 95% CI 0.60 to 0.95; specificity 0.89, 95% CI 0.67 to 0.99) in another study of similar size (41 participants) (Hottat 2009). The latter findings were comparable with those for T1/T2-w MRI with fat-suppression/Gd, which was evaluated in one study comprising 92 participants, which reported sensitivity of 0.84 (95% CI 0.75 to 0.91) and specificity of 0.89 (95% CI 0.52 to 1.00) (Bazot 2009). Overall, although TVUS met the criteria for a SpPin triage test in mapping USL endometriosis, TRUS approached these criteria but

presented wide CIs and insufficient data for meaningful evaluation. MRI displayed the highest sensitivity of all modalities but did not reach SpPin or SnNout criteria.

Direct comparisons of imaging tests for USL endometriosis

- Direct comparison between MRI, TVUS and TRUS performed by Bazot et al. (Bazot 2009) showed that MRI was the most accurate method, and TVUS (sensitivity 0.78, 95% CI 0.68 to 0.87; specificity 0.67, 95% CI 0.30 to 0.93) performed better than TRUS (sensitivity 0.48, 95% CI 0.37 to 0.59; specificity 0.44, 95% CI 0.14 to 0.79) for detection of USL endometriosis (Appendix 5: Figure 29; Figure 30; Figure 31).
- Another direct comparison (23 participants) (Bazot 2013) revealed that 2D-MRI and 3D-MRI had a similar diagnostic performance (sensitivity 0.88, 95% CI 0.64 to 0.99; specificity 0.33, 95% CI 0.04 to 0.78) for both tests (Appendix 5: Figure 32).

Figure 29. Forest plot demonstrating the direct comparison between TVUS and TRUS for USL involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS USL

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	65	3	18	6	No	Europe	0.78 [0.68, 0.87]	0.67 [0.30, 0.93]		

TRUS USL

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	40	5	43	4	No	Europe	0.48 [0.37, 0.59]	0.44 [0.14, 0.79]		

Figure 30. Forest plot demonstrating the direct comparison between MRI and TRUS for USL involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TRUS USL

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	40	5	43	4	No	Europe	0.48 [0.37, 0.59]	0.44 [0.14, 0.79]		

MRI USL

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	70	1	13	8	T1/T2-w + fat-suppressed/Gd	Europe	0.84 [0.75, 0.91]	0.89 [0.52, 1.00]		

Figure 31. Forest plot demonstrating the direct comparison between MRI and TVUS for USL involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

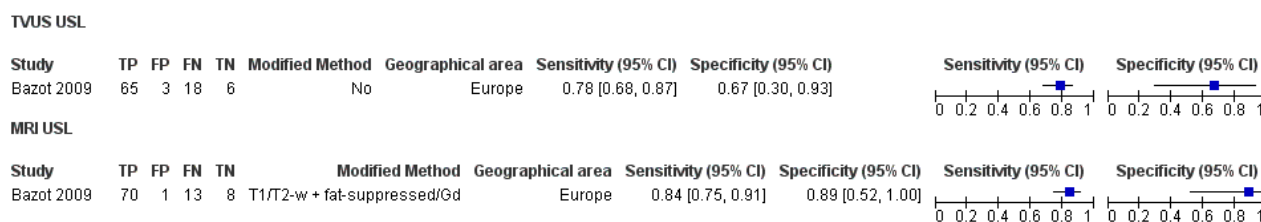
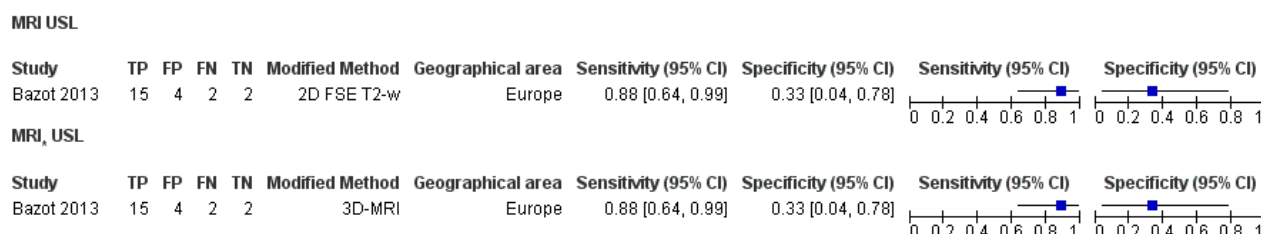


Figure 32. Forest plot demonstrating the direct comparison between 2D-MRI and 3D-MRI for USL involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.



RVS endometriosis

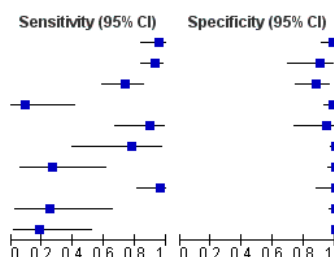
Twelve studies (16 data sets) assessed the diagnostic accuracy of TVUS, TRUS and MRI in detecting RVS endometriosis in Europe ($n = 7$), South America ($n = 3$) and Australia ($n = 2$). For TVUS (10 studies, 11 data sets, 983 participants), mean sensitivity and specificity were 0.88 (95% CI 0.82 to 0.94) and 1.00 (95% CI 0.98 to 1.00), respectively. For MRI (three studies, 288 participants), summary sensitivity and

specificity were 0.81 (95% CI 0.70 to 0.93) and 0.86 (95% CI 0.78 to 0.95), respectively. For TRUS (two studies, 232 participants), summary sensitivity and specificity were 0.78 (95% CI 0.51 to 1.00) and 0.96 (95% CI 0.89 to 1.00), respectively. The heterogeneity of sensitivity was greater than that of specificity for all imaging tests (Figure 33). Substantial scatter of the estimates of sensitivity was evident when TVUS estimates were plotted in the ROC space (Figure 34).

Figure 33. Forest plot of all imaging tests for diagnosis of RVS involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are ordered according to the year of publication for each test. Tests on the same population (different TVUS methods) are presented separately as TVUS*. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

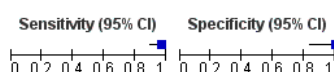
TVUS RVS

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Abrão 2007	39	1	2	62	TVUS-BP	South America	0.95 [0.83, 0.99]	0.98 [0.91, 1.00]
Menada 2008a	64	2	5	19	TVUS-BP	Europe	0.93 [0.84, 0.98]	0.90 [0.70, 0.99]
Guerriero 2008	34	5	12	37	tg-TVUS	Europe	0.74 [0.59, 0.86]	0.88 [0.74, 0.96]
Bazot 2009	1	1	10	80	No	Europe	0.09 [0.00, 0.41]	0.99 [0.93, 1.00]
Pascual 2010	17	1	2	18	3D Introital US	Europe	0.89 [0.67, 0.99]	0.95 [0.74, 1.00]
Hudelist 2011a	7	0	2	120	No	Europe	0.78 [0.40, 0.97]	1.00 [0.97, 1.00]
Falco 2011	3	0	8	85	No	Europe	0.27 [0.06, 0.61]	1.00 [0.96, 1.00]
Scarella 2013	27	0	1	29	TVUS-BP	South America	0.96 [0.82, 1.00]	1.00 [0.88, 1.00]
Reid 2013a	2	0	6	92	No	Australia	0.25 [0.03, 0.65]	1.00 [0.96, 1.00]
Reid 2014	2	0	9	178	SVG	Australia	0.18 [0.02, 0.52]	1.00 [0.98, 1.00]



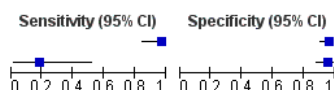
TVUS, RVS

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Menada 2008a	67	0	2	21	RWC-TVS	Europe	0.97 [0.90, 1.00]	1.00 [0.84, 1.00]



TRUS RVS

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Fedele 1998	33	4	1	102	No	Europe	0.97 [0.85, 1.00]	0.96 [0.91, 0.99]
Bazot 2009	2	4	9	77	No	Europe	0.18 [0.02, 0.52]	0.95 [0.88, 0.99]



MRI RVS

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Abrão 2007	31	20	10	43	T1/T2-w + fat-suppressed/Gd	South America	0.76 [0.60, 0.88]	0.68 [0.55, 0.79]
Bazot 2009	6	1	5	80	T1/T2-w + fat-suppressed/Gd	Europe	0.55 [0.23, 0.83]	0.99 [0.93, 1.00]
Chamie 2009a	59	2	7	24	T1/T2-w + fat-suppressed/Gd	South America	0.89 [0.79, 0.96]	0.92 [0.75, 0.99]

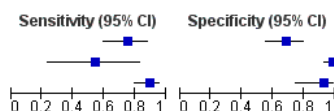
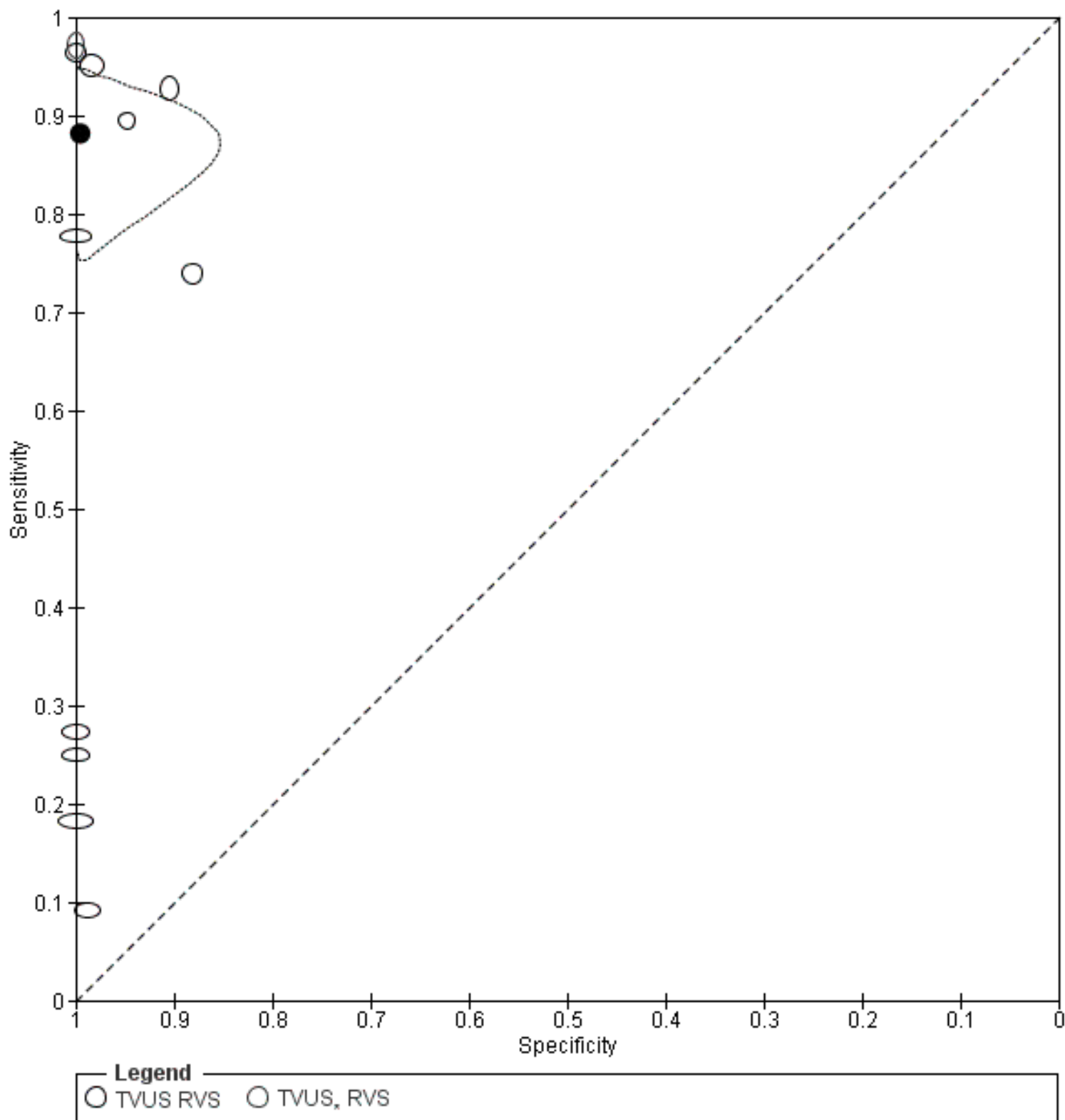


Figure 34. Summary ROC plot of TVUS for detection of RVS involvement. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. Tests on the same population (different TVUS methods) are presented separately as TVUS_s. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).



Indirect comparisons of imaging tests for RVS endometriosis

TVUS-BP studies (three studies, 250 participants) (Abrao 2007; Menada 2008a; Scarella 2013) and RWC-TVS (one study, 90 participants) (Menada 2008a) demonstrated the highest diagnostic accuracy, with sensitivities ranging from 0.93 to 0.97 and specificities ranging from 0.90 to 1.00. Both TVUS and TRUS met the criteria for a SpPin triage test. TRUS could not be adequately assessed because of the paucity of data and displayed lower diagnostic estimates and wider CIs compared with TVUS. MRI did not meet the criteria for either of the triage tests, but data were insufficient for assessment of its role in a meaningful way.

Direct comparisons of imaging tests for RVS endometriosis

- Direct comparison (one article, 90 participants) (Menada 2008a) showed that TVUS (RWC-TVS) (sensitivity 0.97, 95% CI 0.90 to 1.00; specificity 1.00, 95% CI 0.84 to 1.00) displayed greater accuracy than conventional TVUS (sensitivity 0.93, 95% CI 0.84

to 0.98; specificity 0.90, 95% CI 0.70 to 0.99) in detecting RVS endometriosis (Appendix 6: Figure 35).

- When TRUS and TVUS were directly compared (one study, 92 participants) (Bazot 2009), sensitivities were very low for both (0.18, 95% CI 0.02 to 0.52; 0.09, 95% CI 0.00 to 0.41), respectively, although TVUS had higher specificity (0.99, 95% CI 0.93 to 1.00) than TRUS (0.95, 95% CI 0.88 to 0.99) (Appendix 6: Figure 36). The same study revealed that TRUS and TVUS appeared to be less sensitive than MRI (sensitivity 0.55, 95% CI 0.23 to 0.83; specificity 0.99, 95% CI 0.93 to 1.00); specificity for MRI was higher than for TRUS and comparable with that for TVUS (Appendix 6: Figure 37; Figure 38).
- In contrast, another comparative study of 104 participants (Abrao 2007) showed that TVUS (sensitivity 0.95, 95% CI 0.83 to 0.99; specificity 0.98, 95% CI 0.91 to 1.00) yielded higher diagnostic estimates than MRI (sensitivity of 0.76, 95% CI 0.60 to 0.88; specificity 0.68, 95% CI 0.55 to 0.79) for detection of RVS endometriosis (Appendix 6: Figure 38).

Figure 35. Forest plot demonstrating the direct comparison between TVUS and RWC-TVS for RVS involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS RVS

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Menada 2008a	64	2	5	19	TVUS-BP	Europe	0.93 [0.84, 0.98]	0.90 [0.70, 0.99]		

TVUS₁ RVS

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Menada 2008a	67	0	2	21	RWC-TVS	Europe	0.97 [0.90, 1.00]	1.00 [0.84, 1.00]		

Figure 36. Forest plot demonstrating the direct comparison between TVUS and TRUS for RVS involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS RVS

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	1	1	10	80	No	Europe	0.09 [0.00, 0.41]	0.99 [0.93, 1.00]		

TRUS RVS

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	2	4	9	77	No	Europe	0.18 [0.02, 0.52]	0.95 [0.88, 0.99]		

Figure 37. Forest plot demonstrating the direct comparison between MRI and TRUS for RVS involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

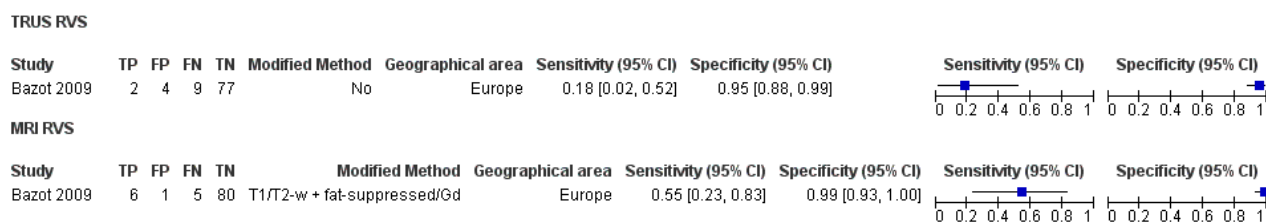
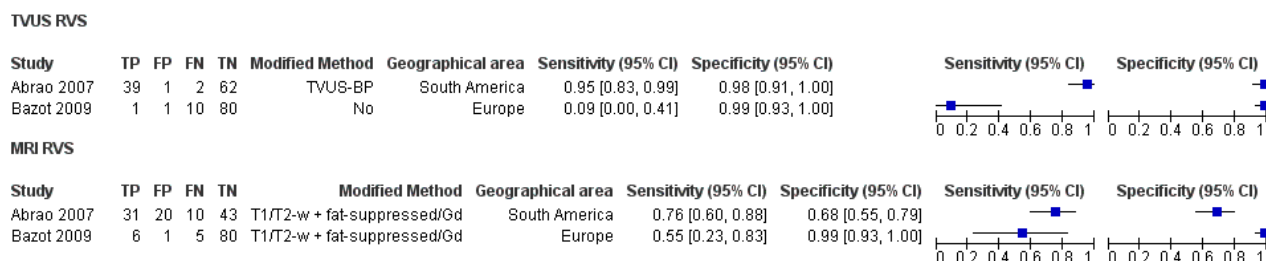


Figure 38. Forest plot demonstrating the direct comparison between MRI and TVUS for RVS involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.



The confidence intervals were wide and overlapping in all direct comparisons, and data were insufficient data for statistical comparison of the different imaging modalities for RVS endometriosis.

Vaginal wall endometriosis

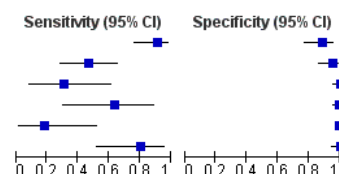
Ten studies (13 data sets) assessed the diagnostic accuracy of TVUS, TRUS and MRI for detecting vaginal wall endometriosis in Europe (n = 7), South America (n = 1) and Australia (n = 2). For TVUS (six studies, 679 participants), mean sensitivity and mean specificity

were 0.57 (95% CI 0.21 to 0.94) and 0.99 (95% CI 0.96 to 1.00), respectively. For MRI (four studies, five data sets, 248 participants), mean sensitivity and specificity were 0.77 (95% CI 0.67 to 0.88) and 0.97 (95% CI 0.92 to 1.00), respectively. In the two studies that evaluated TRUS in 232 participants, summary sensitivity and specificity were 0.39 (95% CI 0.08 to 0.70) and 1.00 (95% CI 1.00 to 1.00), respectively. Heterogeneity was greater for estimates of sensitivity than specificity for all test modalities (Figure 39; Figure 40; Figure 41).

Figure 39. Forest plot of all imaging tests for diagnosis of vaginal wall involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are ordered according to year of publication for each test. Tests on the same population (different MRI methods) are presented separately as MRI+, FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

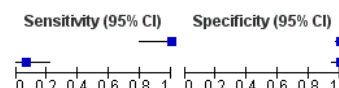
TVUS vaginal

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Guemiero 2008	31	6	3	48	tg-TVUS	Europe	0.91 [0.76, 0.98]	0.89 [0.77, 0.96]
Bazot 2009	14	3	16	59	No	Europe	0.47 [0.28, 0.66]	0.95 [0.87, 0.99]
Falco 2011	4	0	9	83	No	Europe	0.31 [0.09, 0.61]	1.00 [0.96, 1.00]
Hudelist 2011a	7	1	4	117	No	Europe	0.64 [0.31, 0.89]	0.99 [0.95, 1.00]
Reid 2014	2	1	9	177	SVG	Australia	0.18 [0.02, 0.52]	0.99 [0.97, 1.00]
Piessens 2014	12	0	3	70	TVUS-BP	Australia	0.80 [0.52, 0.96]	1.00 [0.95, 1.00]



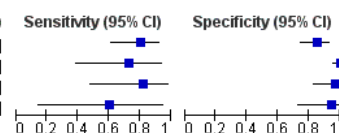
TRUS vaginal

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Fedele 1998	16	0	0	124	No	Europe	1.00 [0.79, 1.00]	1.00 [0.97, 1.00]
Bazot 2009	2	0	28	62	No	Europe	0.07 [0.01, 0.22]	1.00 [0.94, 1.00]



MRI vaginal

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	24	9	6	53	T1/T2-w + fat-suppressed/Gd	Europe	0.80 [0.61, 0.92]	0.85 [0.74, 0.93]
Chamie 2009a	8	0	3	81	T1/T2-w + fat-suppressed/Gd	South America	0.73 [0.39, 0.94]	1.00 [0.96, 1.00]
Hottat 2009	9	1	2	29	3.0T	Europe	0.82 [0.48, 0.98]	0.97 [0.83, 1.00]
Bazot 2013	3	1	2	17	2D FSE T2-w	Europe	0.60 [0.15, 0.95]	0.94 [0.73, 1.00]



MRI, vaginal

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2013	4	0	1	18	3D-MRI	Europe	0.80 [0.28, 0.99]	1.00 [0.81, 1.00]

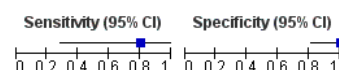


Figure 40. Summary ROC plot of TVUS for detection of vaginal wall involvement. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).

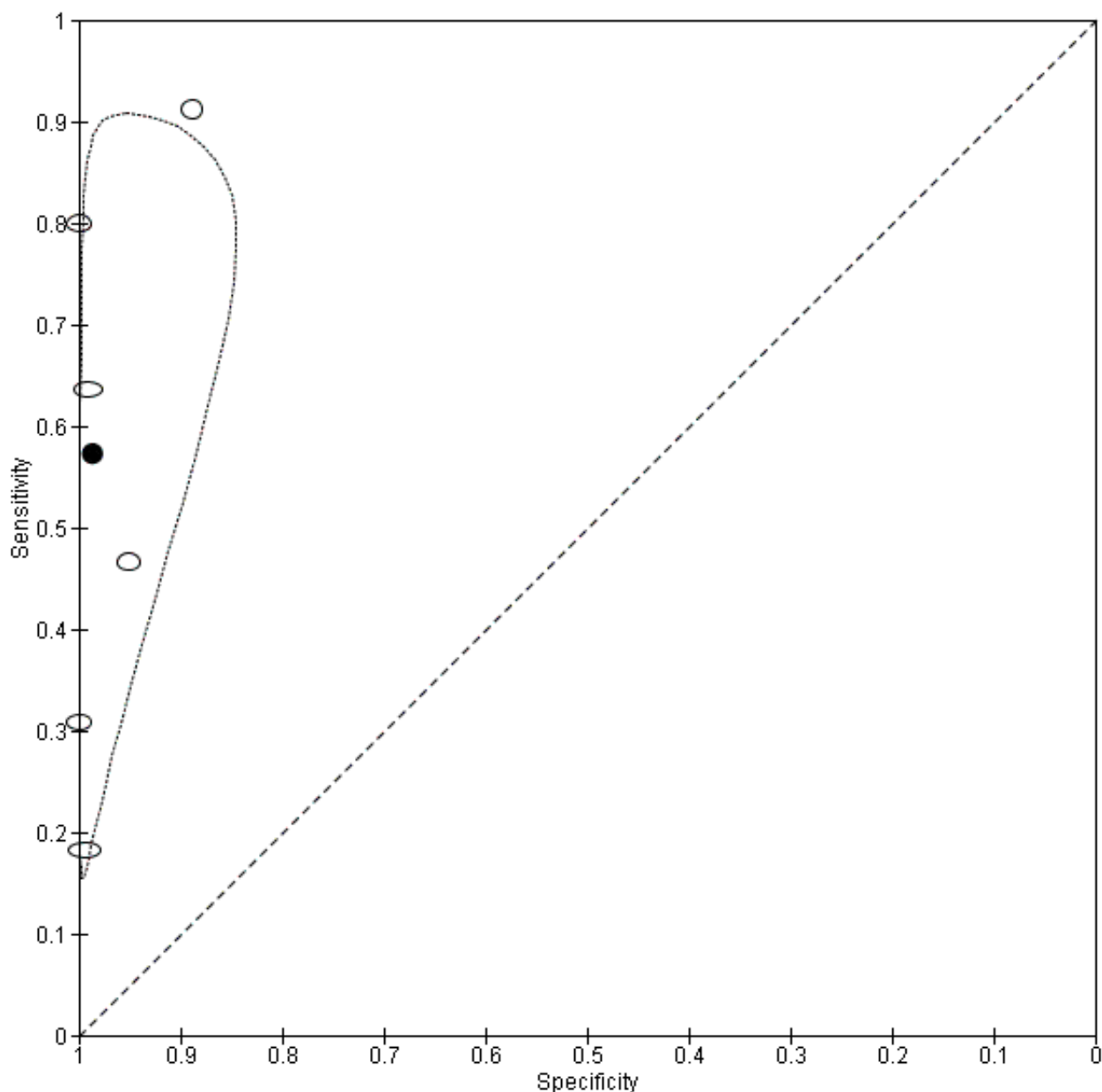
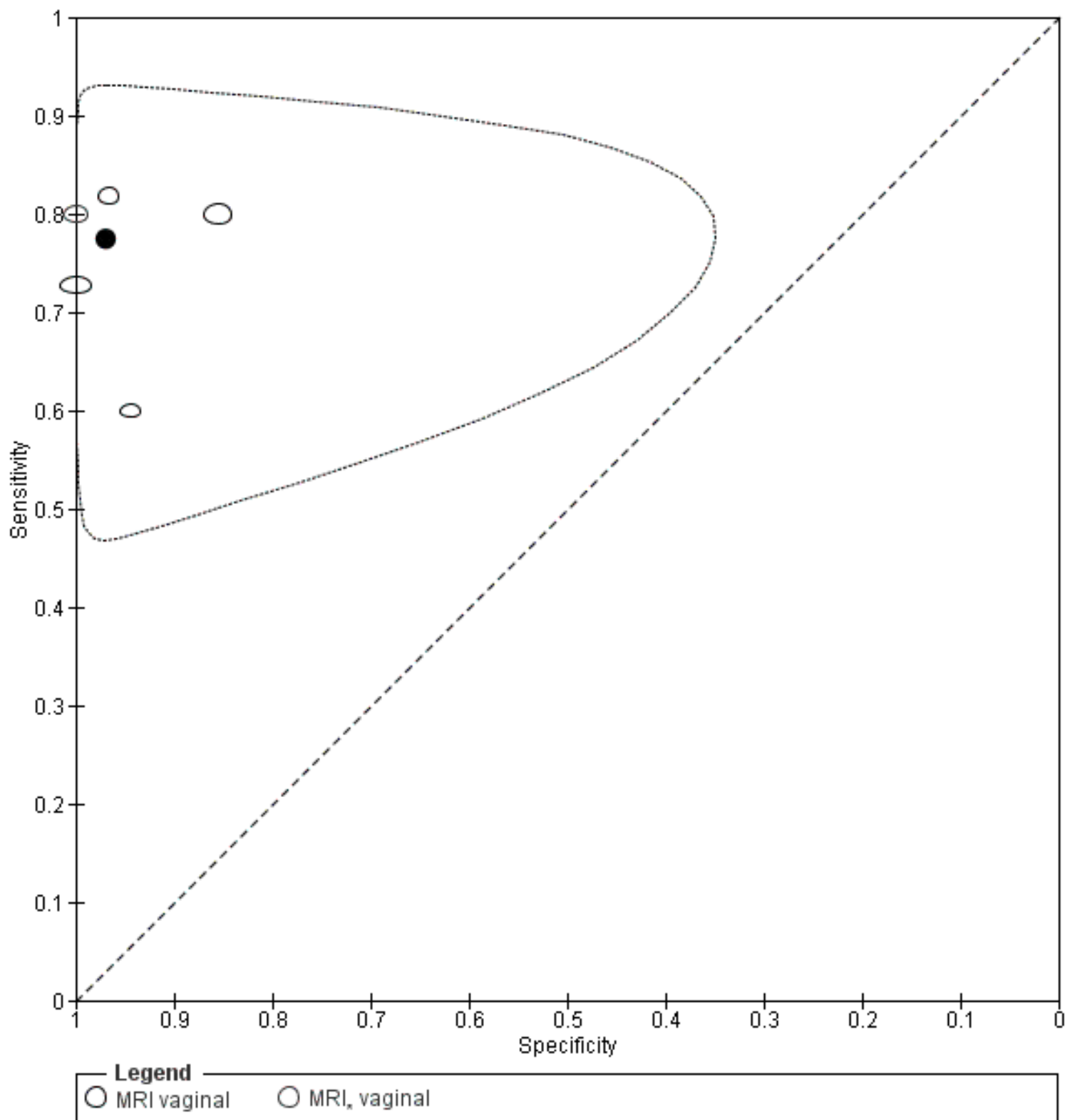


Figure 41. Summary ROC plot of MRI for detection of vaginal wall involvement. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. Tests on the same population (different MRI methods) are presented separately as MRI*. The solid black circle represents the mean sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).



Indirect comparisons of imaging tests for vaginal wall endometriosis

Tg-TVUS (one study, 88 participants) had the highest diagnostic estimates among TVUS methods (sensitivity 0.91, 95% CI 0.76 to 0.98; specificity 0.89, 95% CI 0.77 to 0.96) (Guerriero 2008). 3D MRI (one study, 23 participants) (Bazot 2013) and 3.0T MRI (one study, 41 participants) (Hottat 2009) were the best performing MRI modalities with sensitivities of 0.80 and 0.82 (95% CI 0.28 to 0.99 and 0.48 to 0.98) and specificities of 1.0 and 0.97 (95% CI 0.81 to 1.00 and 0.83 to 1.00), respectively. Both TVUS and MRI met the criteria for a SpPin triage test. TVUS showed lower sensitivity but higher specificity compared with MRI. For TRUS, the criteria for either triage test were not met and CIs were wide, although data were insufficient data to permit meaningful conclusions.

Direct comparisons of imaging tests for vaginal wall endometriosis

- In a direct comparison comprising 92 participants (Bazot 2009), MRI (sensitivity 0.80, 95% CI 0.61 to 0.92; specificity, 0.86, 95% CI 0.74 to 0.93) showed higher sensitivity but lower specificity than TVUS (sensitivity 0.47, 95% CI 0.28 to 0.66; specificity 0.95, 95% CI 0.87 to 0.99) and TRUS (sensitivity 0.07, 95% CI 0.10 to 0.22; specificity 1.00, 95% CI 0.94 to 1.00); TRUS had much lower sensitivity but higher specificity than either TVUS or MRI (Appendix 7: Figure 42; Figure 43; Figure 44).
- 2D-MRI (sensitivity 0.60, 95% CI 0.15 to 0.95; specificity 0.94, 95% CI 0.73 to 1.00) demonstrated lower accuracy estimates than 3D-MRI (sensitivity 0.8, 95% CI 0.28 to 0.99; specificity 1.00, 95% CI 0.81 to 1.00) in a paired comparative study of 23 participants (Bazot 2013) (Appendix 7: Figure 45).

Figure 42. Forest plot demonstrating the direct comparison between TVUS and TRUS for vaginal wall involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

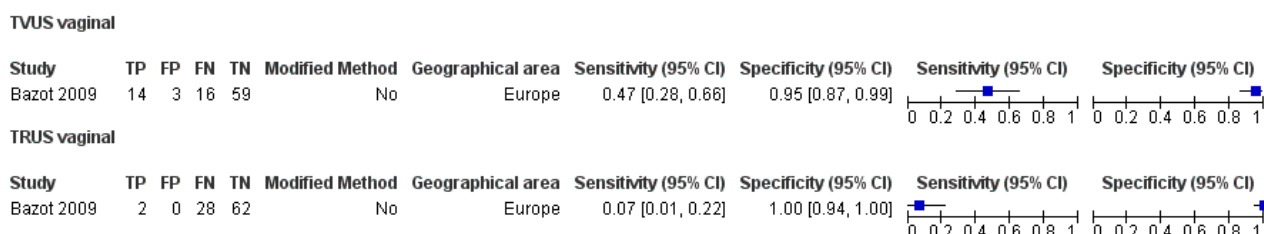


Figure 43. Forest plot demonstrating the direct comparison between TRUS and MRI for vaginal wall involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

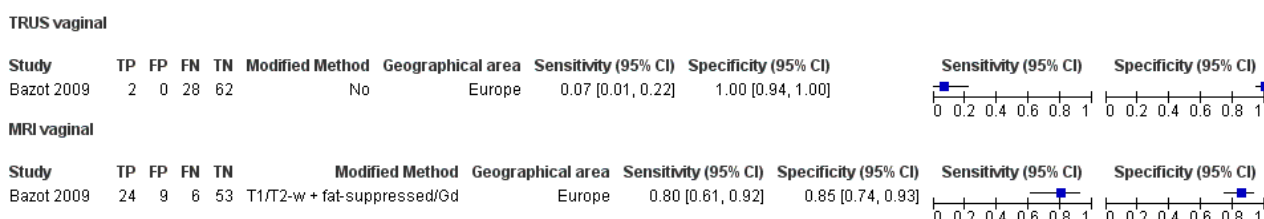


Figure 44. Forest plot demonstrating the direct comparison between TVUS and MRI for vaginal wall involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

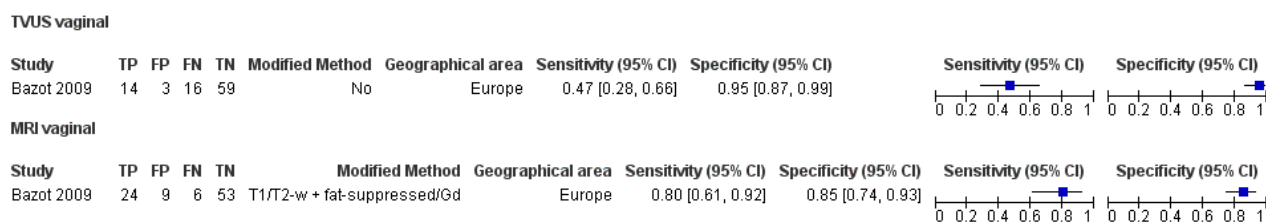
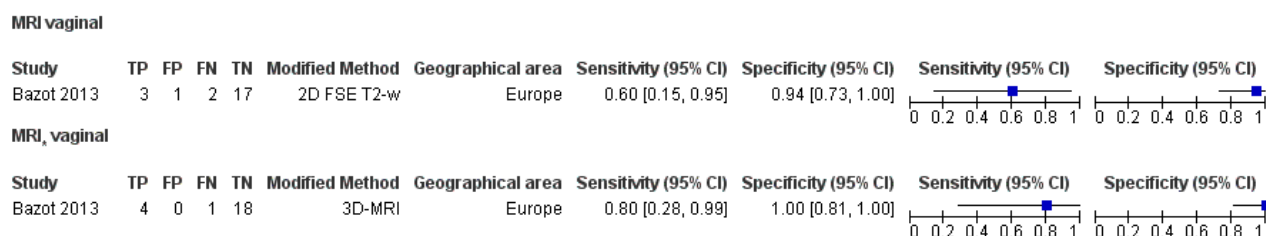


Figure 45. Forest plot demonstrating the direct comparison between 2D-MRI and 3D-MRI for vaginal wall involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.



POD obliteration by endometriosis

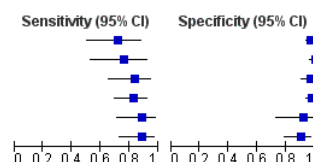
Eleven publications (12 data sets) assessed the diagnostic accuracy of TVUS and MRI for detecting an obliterated POD in endometriosis in Europe ($n = 6$), Australia ($n = 3$), South America ($n = 1$) and Asia ($n = 1$). For TVUS (six studies, 755 participants), mean sensitivity and specificity were 0.83 (95% CI 0.77, 0.88) and 0.97 (95% CI

0.95 to 0.99), respectively. For MRI (five studies, six data sets, 154 participants), mean sensitivity and specificity were 0.90 (95% CI 0.76 to 1.00) and 0.98 (95% CI 0.89 to 1.00), respectively. Heterogeneity was greater for sensitivity than for specificity for TVUS, whereas both estimates were heterogeneous for MRI (Figure 46; Figure 47; Figure 48).

Figure 46. Forest plot of all imaging tests for diagnosis of POD obliteration by endometriosis. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are ordered according to year of publication for each test. Tests on the same population (different MRI methods) are presented separately as MRI*. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

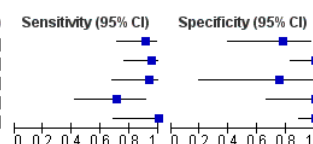
TVUS POD

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Holland 2010	18	5	7	171	No	Europe	0.72 [0.51, 0.88]	0.97 [0.93, 0.99]
Hudellist 2011a	16	0	5	108	No	Europe	0.76 [0.53, 0.92]	1.00 [0.97, 1.00]
Reid 2013a	25	2	5	68	No	Australia	0.83 [0.65, 0.94]	0.97 [0.90, 1.00]
Reid 2014	39	3	8	139	No	Australia	0.83 [0.69, 0.92]	0.98 [0.94, 1.00]
Leon 2014	24	2	3	22	SVG + TVUS-BP	South America	0.89 [0.71, 0.98]	0.92 [0.73, 0.99]
Piessens 2014	30	5	4	46	TVUS-BP	Australia	0.88 [0.73, 0.97]	0.90 [0.79, 0.97]



MRI POD

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Takeuchi 2005	20	2	2	7	Jelly method [T1/T2-w + fat-suppressed]	Asia	0.91 [0.71, 0.99]	0.78 [0.40, 0.97]
Hottat 2009	20	0	1	20	3.0T	Europe	0.95 [0.76, 1.00]	1.00 [0.83, 1.00]
Manganaro 2012b	14	1	1	3	3.0T	Europe	0.93 [0.68, 1.00]	0.75 [0.19, 0.99]
Bazot 2013	10	0	4	9	2D FSE T2-w	Europe	0.71 [0.42, 0.92]	1.00 [0.66, 1.00]
Thomeer 2014	10	0	0	30	3.0T	Europe	1.00 [0.69, 1.00]	1.00 [0.88, 1.00]



MRI, POD

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2013	10	0	4	9	3D-MRI	Europe	0.71 [0.42, 0.92]	1.00 [0.66, 1.00]

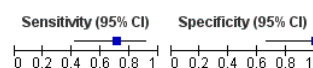


Figure 47. Summary ROC plot of TVUS for detection of POD obliteration by endometriosis. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).

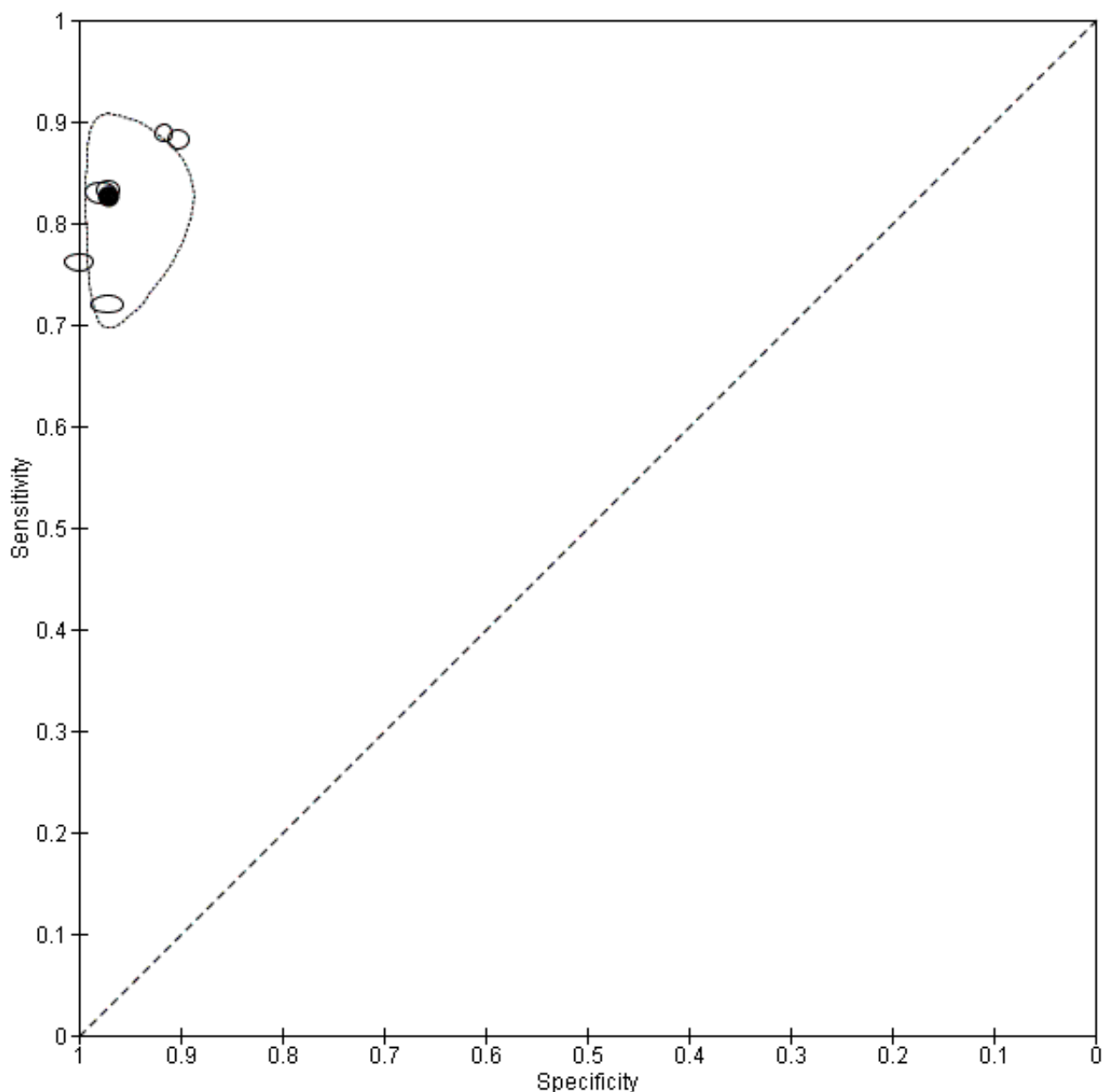
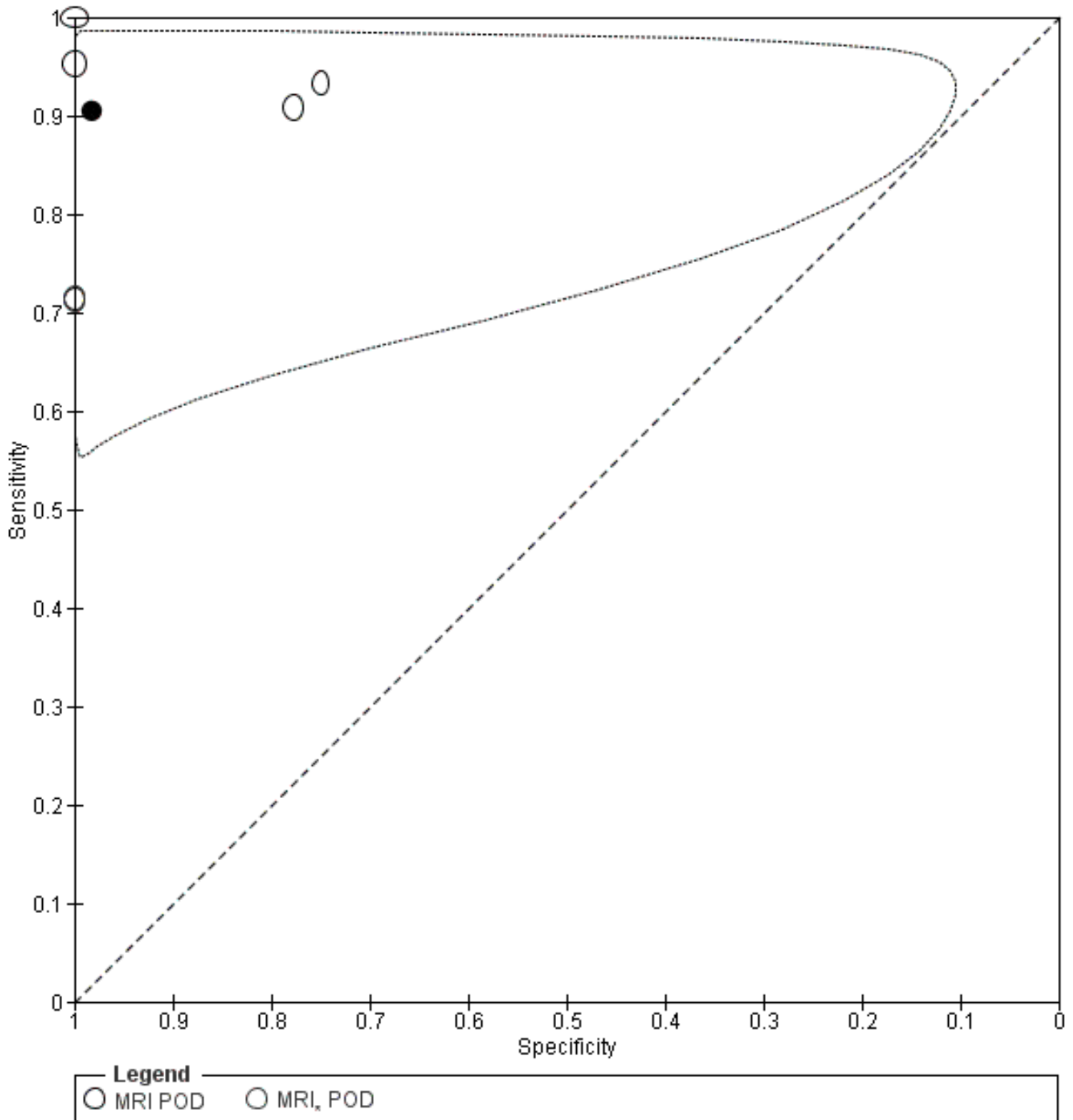


Figure 48. Summary ROC plot of MRI for detection of POD obliteration by endometriosis. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. Tests on the same population (different MRI methods) are presented separately as MRI*. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).



Indirect comparisons of imaging tests for POD obliteration by endometriosis

TVUS-BP (two studies, 136 participants) demonstrated the highest diagnostic accuracy of all TVUS methods with sensitivities of 0.89 and 0.88 (95% CI 0.71 to 0.98 and 0.73 to 0.97) and specificities of 0.92 and 0.90 (95% CI 0.73 to 0.99 and 0.79 to 0.97) (Leon 2014; Piessens 2014). 3.0T MRI (three studies, 100 participants) was the best performing MRI technique with sensitivities ranging from 0.93 to 1.00 and specificities ranging from 0.75 to 1.00 (Hottat 2009; Mangano 2012a; Thomeer 2014). Both TVUS and MRI could qualify as a SpPin triage test for detecting POD obliteration in

endometriosis with slightly higher diagnostic estimates for MRI, which also approached the criteria for a SnNout triage test.

Direct comparisons of imaging tests for POD obliteration by endometriosis

2D-MRI had similar accuracy to 3D-MRI for detection of POD obliteration with sensitivity of 0.71 (95% CI 0.42 to 0.92) and specificity of 1.00 (95% CI 0.66 to 1.00) for both data sets in one small direct comparison comprising 23 participants (Bazot 2013) (Appendix 7: Figure 49).

Figure 49. Forest plot demonstrating the direct comparison between 2D-MRI and 3D-MRI for POD obliteration by endometriosis. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

MRI POD

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2013	10	0	4	9	2D FSE T2-w	Europe	0.71 [0.42, 0.92]	1.00 [0.66, 1.00]		

MRI, POD

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2013	10	0	4	9	3D-MRI	Europe	0.71 [0.42, 0.92]	1.00 [0.66, 1.00]		

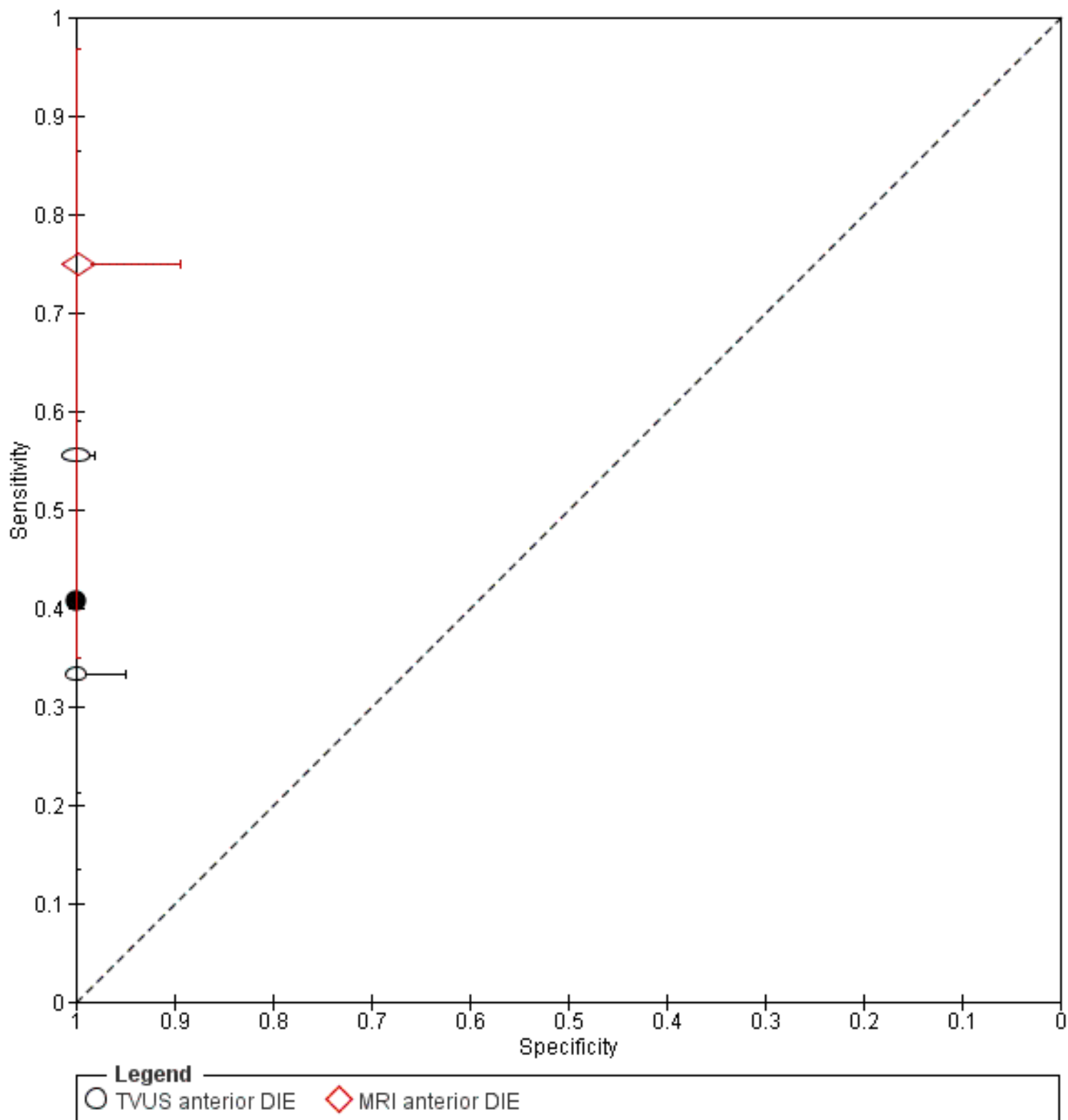
Anterior DIE

Three studies assessed the diagnostic accuracy of TVUS and MRI in diagnosing anterior DIE in Europe. For TVUS (two studies, 289 participants), summary sensitivity and specificity were 0.41 (95% CI 0.00 to 0.81) and 1.00 (95% CI 1.00 to 1.00). MRI (one study,

41 participants) demonstrated sensitivity of 0.75 (95% CI 0.35 to 0.97) and specificity of 1.00 (95% CI 0.89 to 1.00) in detecting anterior DIE (Figure 50). The diagnostic accuracy of bladder and ureteric endometriosis was not assessed in this review (see Target conditions).

Figure 50. Summary ROC plot of TVUS and MRI for detection of anterior DIE. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size, and

the shape designates different imaging modalities. The solid black circle represents the pooled sensitivity and specificity for TVUS, and the bars correspond to 95% CIs of each individual study.



Rectosigmoid endometriosis

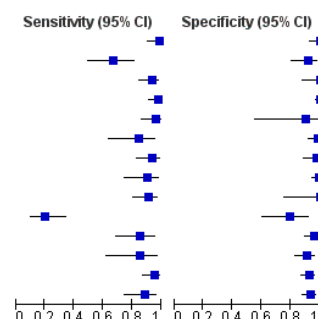
A total of 21 studies (31 data sets) assessed the accuracy of TVUS, TRUS, MRI, MDCT-e and DCBE for detecting rectosigmoid endometriosis in Europe (n = 15), South America (n = 4) and Australia (n = 2). Mean estimates for each imaging modality were as follows: for TVUS (14 studies, 15 data sets, 1616 participants), sensitivity of 0.90 (95% CI 0.82 to 0.97) and specificity of 0.96 (95% CI 0.94 to 0.99); for TRUS (four studies, 330 participants), sensitivity of 0.91 (95% CI 0.85 to 0.98) and specificity of 0.96 (95% CI 0.91 to 1.00); for MRI (six studies, seven data sets, 612 participants),

sensitivity of 0.92 (95% CI 0.86 to 0.99) and specificity of 0.96 (95% CI 0.93 to 0.98). Less heterogeneity was seen in the estimates for TVUS, TRUS and MRI in rectosigmoid endometriosis than in other anatomical locations (Figure 51; Figure 52; Figure 53; Figure 54). For MDCT-e (three studies, 389 participants), summary sensitivity and specificity were 0.98 (95% CI 0.94 to 1.00) and 0.99 (95% CI 0.97 to 1.00) (Figure 55). For DCBE (two studies, 106 participants), summary sensitivity and specificity were 0.56 (95% CI 0.32 to 0.80) and 0.77 (95% CI 0.41 to 1.00), and both studies displayed considerable heterogeneity (Figure 56).

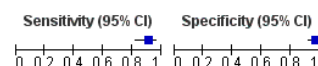
Figure 51. Forest plot of all imaging tests for diagnosis of rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are ordered according to year of publication for each test. Tests on the same population (different TVUS and MRI methods) are presented separately as TVUS* and MRI*. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS rectosigmoid

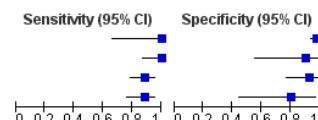
Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Abrao 2007	53	0	1	50	TVUS-BP	South America	0.98 [0.90, 1.00]	1.00 [0.93, 1.00]
Guerriero 2008	26	4	13	45	tg-TVUS	Europe	0.67 [0.50, 0.81]	0.92 [0.80, 0.98]
Bazot 2009	59	0	4	29	No	Europe	0.94 [0.85, 0.98]	1.00 [0.88, 1.00]
Goncalves 2010	79	0	2	113	TVUS-BP	South America	0.98 [0.91, 1.00]	1.00 [0.97, 1.00]
Bergamini 2010	49	1	2	9	RWC-TVUS	Europe	0.96 [0.87, 1.00]	0.90 [0.55, 1.00]
Falco 2011	21	1	4	70	No	Europe	0.84 [0.64, 0.95]	0.99 [0.92, 1.00]
Ferrero 2011	45	1	3	47	RWC-TVUS	Europe	0.94 [0.83, 0.99]	0.98 [0.89, 1.00]
Hudelist 2011a	28	1	3	97	No	Europe	0.90 [0.74, 0.98]	0.99 [0.94, 1.00]
Savelli 2011	51	0	5	13	No	Europe	0.91 [0.80, 0.97]	1.00 [0.75, 1.00]
Mangler 2013	10	6	39	23	No	Europe	0.20 [0.10, 0.34]	0.79 [0.60, 0.92]
Hudelist 2013	29	3	5	80	No	Europe	0.85 [0.69, 0.95]	0.96 [0.90, 0.99]
Reid 2013a	17	7	3	73	No	Australia	0.85 [0.62, 0.97]	0.91 [0.83, 0.96]
Guerriero 2014	73	9	4	116	tg-TVUS	Europe	0.95 [0.87, 0.99]	0.93 [0.87, 0.97]
Reid 2014	38	10	5	136	SVG	Australia	0.88 [0.75, 0.96]	0.93 [0.88, 0.97]

**TVUS, rectosigmoid**

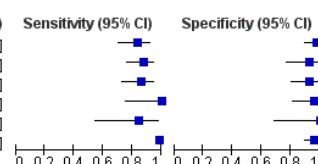
Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 2014	70	4	7	121	3D-TVUS	Europe	0.91 [0.82, 0.96]	0.97 [0.92, 0.99]

**TRUS rectosigmoid**

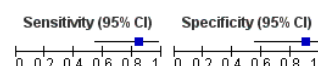
Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Fedele 1998	9	3	0	128	No	Europe	1.00 [0.66, 1.00]	0.98 [0.93, 1.00]
Ribeiro 2008a	27	1	0	9	No	South America	1.00 [0.87, 1.00]	0.90 [0.55, 1.00]
Bazot 2009	56	2	7	27	No	Europe	0.89 [0.78, 0.95]	0.93 [0.77, 0.99]
Bergamini 2010	45	2	6	8	No	Europe	0.88 [0.76, 0.96]	0.80 [0.44, 0.97]

**MRI rectosigmoid**

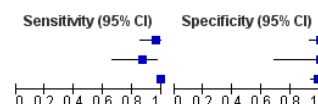
Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Abrao 2007	45	1	9	49	T1/T2-w + fat-suppressed/Gd	South America	0.83 [0.71, 0.92]	0.98 [0.89, 1.00]
Bazot 2009	55	2	8	27	T1/T2-w + fat-suppressed/Gd	Europe	0.87 [0.77, 0.94]	0.93 [0.77, 0.99]
Chamie 2009a	43	3	7	39	T1/T2-w + fat-suppressed/Gd	South America	0.86 [0.73, 0.94]	0.93 [0.81, 0.99]
Hottat 2009	13	1	0	27	3.0T	Europe	1.00 [0.75, 1.00]	0.96 [0.82, 1.00]
Bazot 2013	11	0	2	10	2D FSE T2-w	Europe	0.85 [0.55, 0.98]	1.00 [0.69, 1.00]
Biscaldi 2014	174	3	2	81	Jelly method [T1/T2-w + fat-suppressed]	Europe	0.99 [0.96, 1.00]	0.96 [0.90, 0.99]

**MRI, rectosigmoid**

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2013	11	1	2	9	3D-MRI	Europe	0.85 [0.55, 0.98]	0.90 [0.55, 1.00]

**MDCT-e rectosigmoid**

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Ferrero 2011	46	0	2	48	No	Europe	0.96 [0.86, 0.99]	1.00 [0.93, 1.00]
Stabile 2013	20	0	3	10	No	Europe	0.87 [0.66, 0.97]	1.00 [0.69, 1.00]
Biscaldi 2014	175	1	1	83	No	Europe	0.99 [0.97, 1.00]	0.99 [0.94, 1.00]

**DCBE rectosigmoid**

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)
Ribeiro 2008a	21	6	3	7	No	South America	0.88 [0.68, 0.97]	0.54 [0.25, 0.81]
Savelli 2011	24	0	32	13	No	Europe	0.43 [0.30, 0.57]	1.00 [0.75, 1.00]

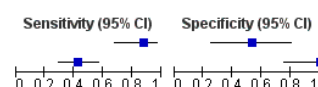


Figure 52. Summary ROC plot of TVUS for detection of rectosigmoid involvement. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. Tests on the same population (different TVUS methods) are presented separately as TVUS_s. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).

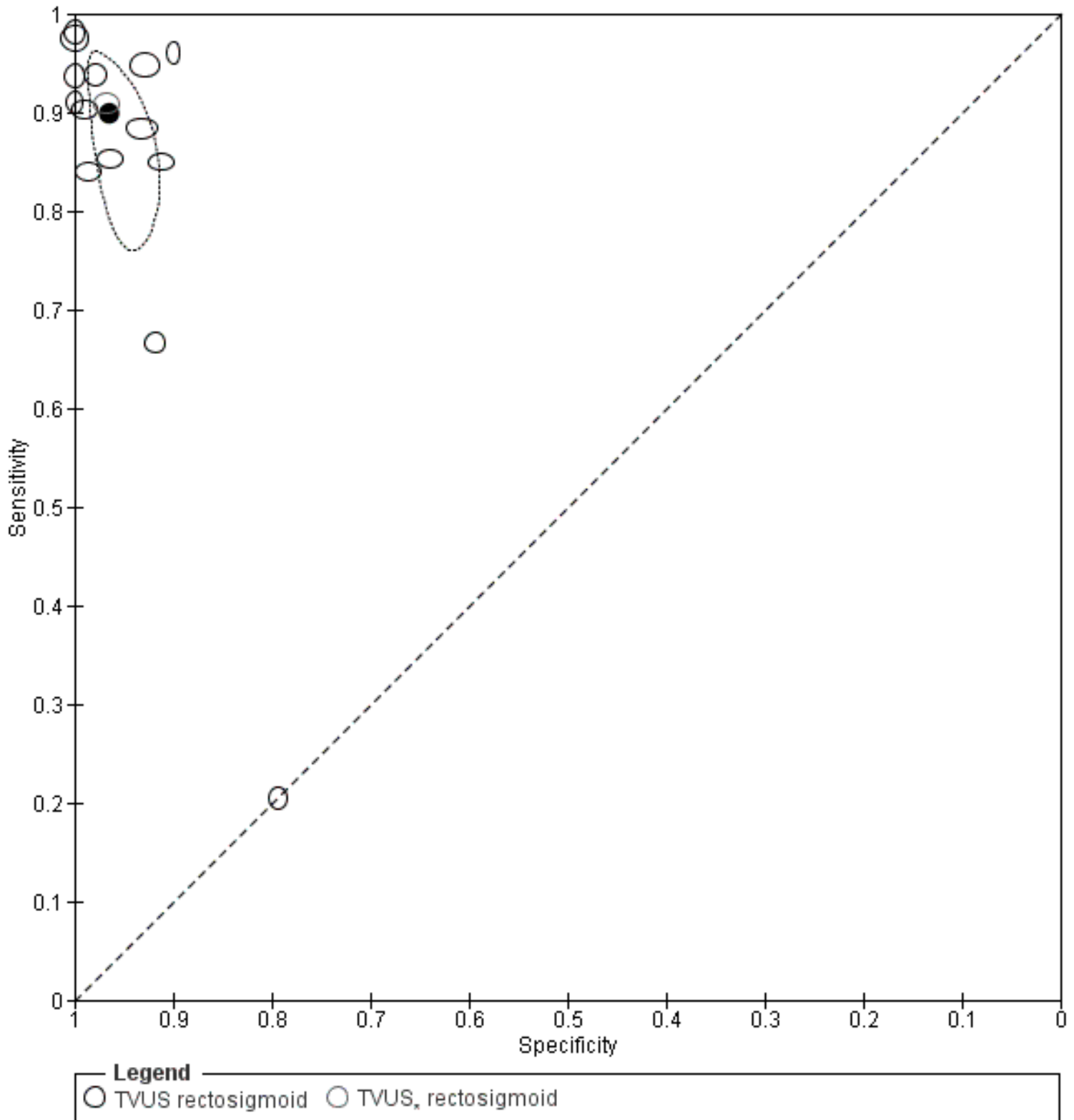


Figure 53. Summary ROC plot of TRUS for detection of rectosigmoid involvement. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).

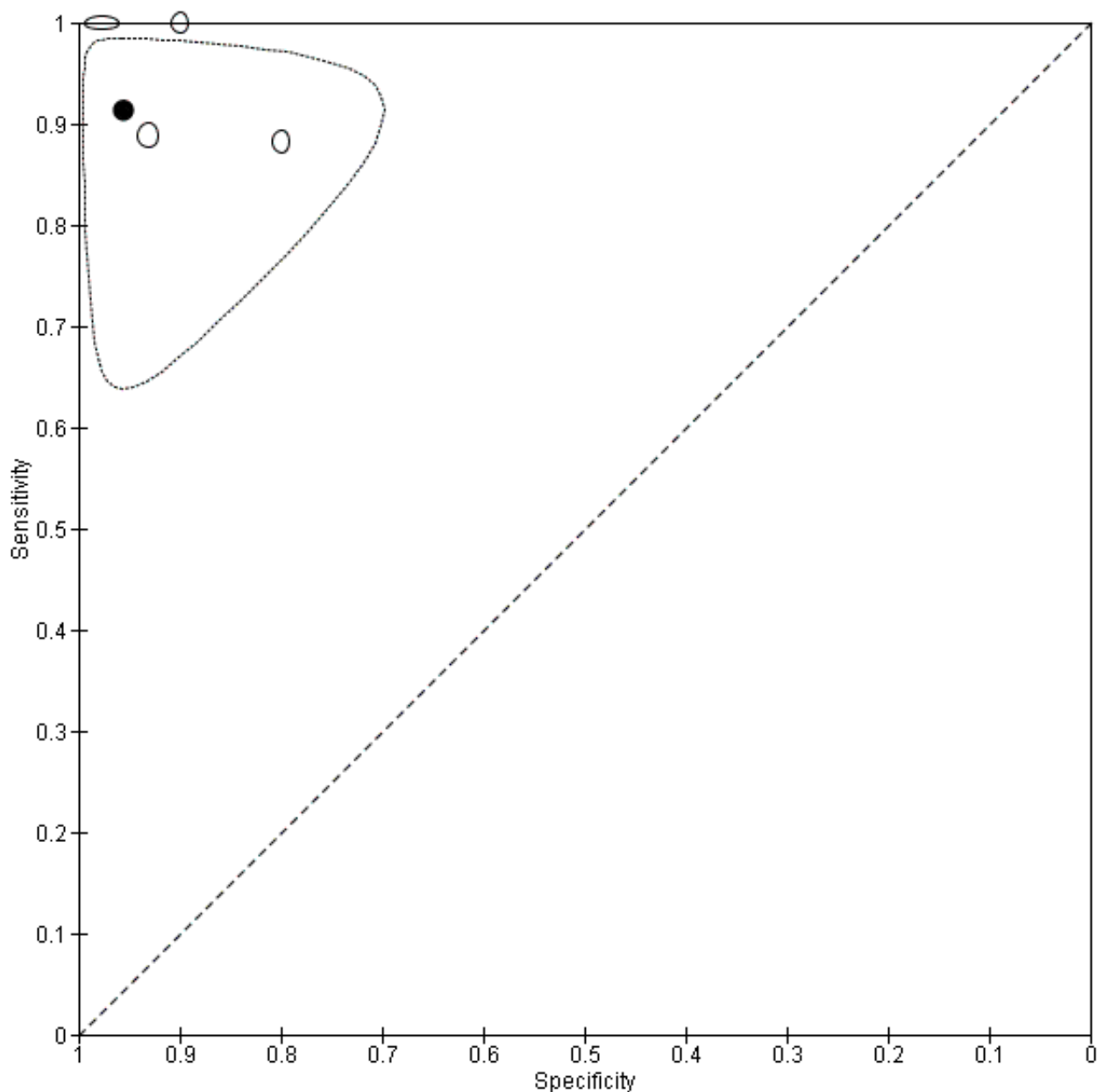


Figure 54. Summary ROC plot of MRI for detection of rectosigmoid involvement. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. Tests on the same population (different MRI methods) are presented separately as MRI*. The solid black circle represents the pooled sensitivity and specificity, which is surrounded by a 95% confidence region (dashed line).

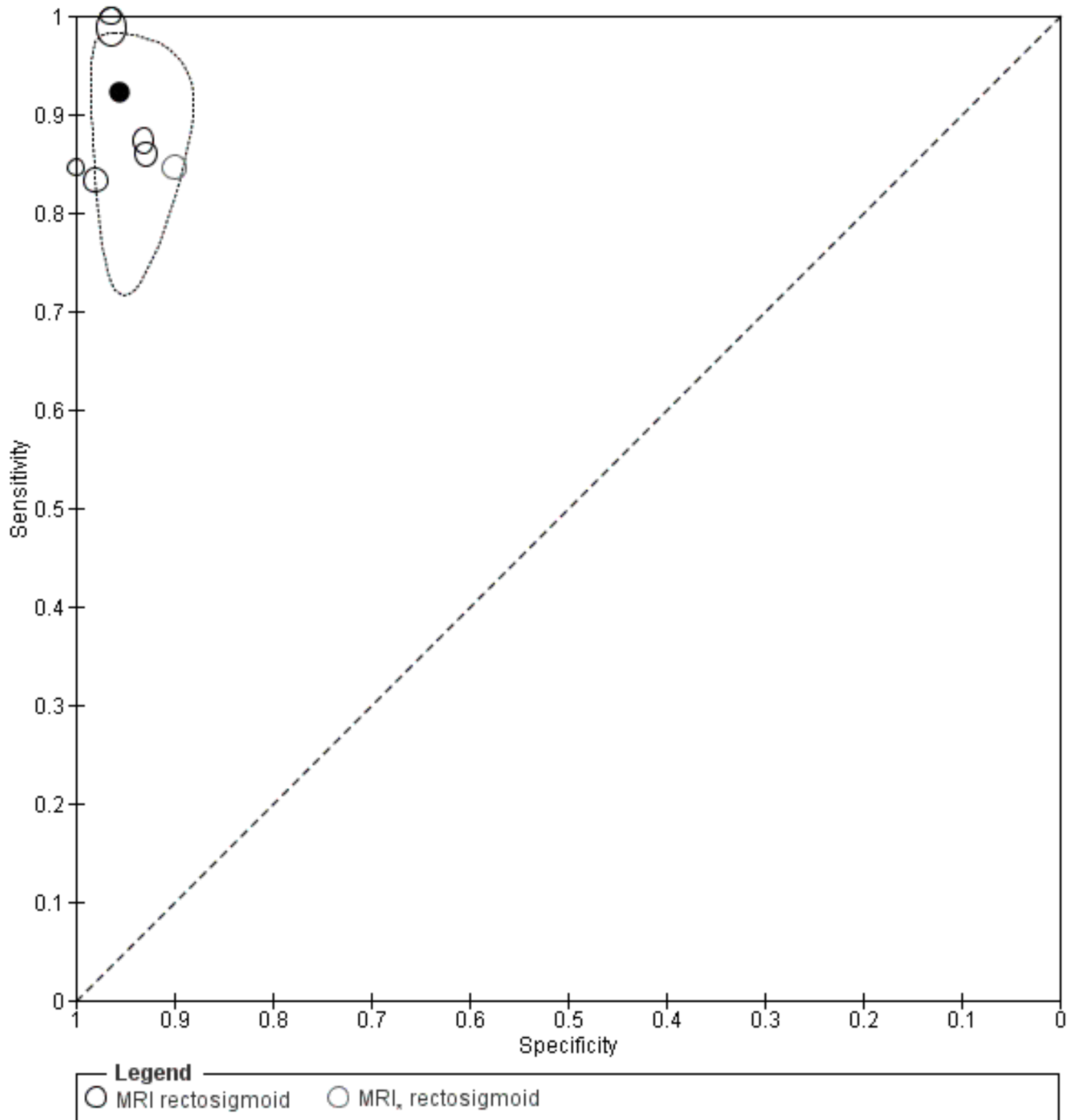


Figure 55. Summary ROC plot of MDCT-e for detection of rectosigmoid involvement. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size and the shape designates consecutive or non-consecutive enrolment. The solid black circle represents the pooled sensitivity and specificity, and the bars correspond to 95% CIs of each individual study.

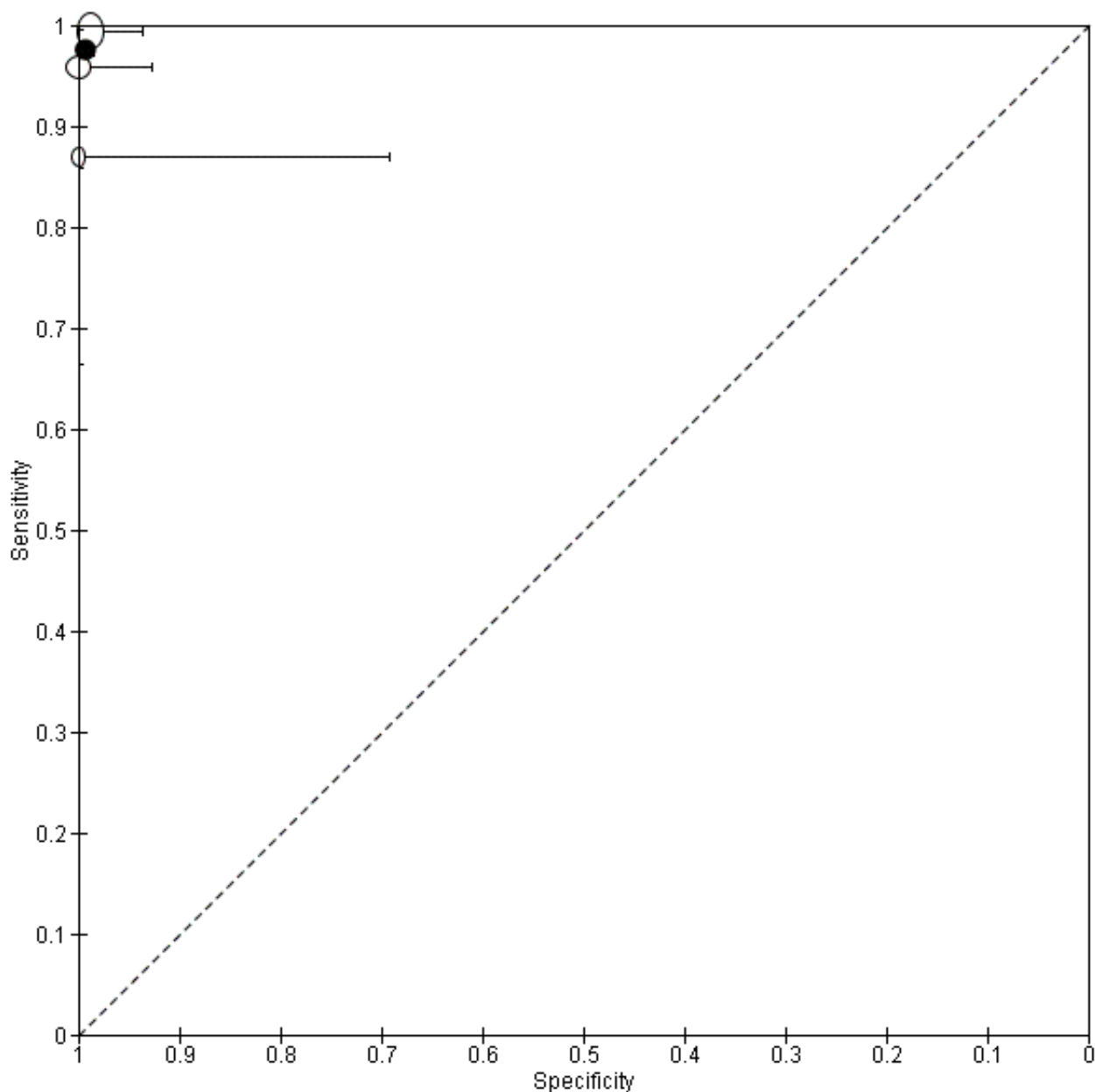
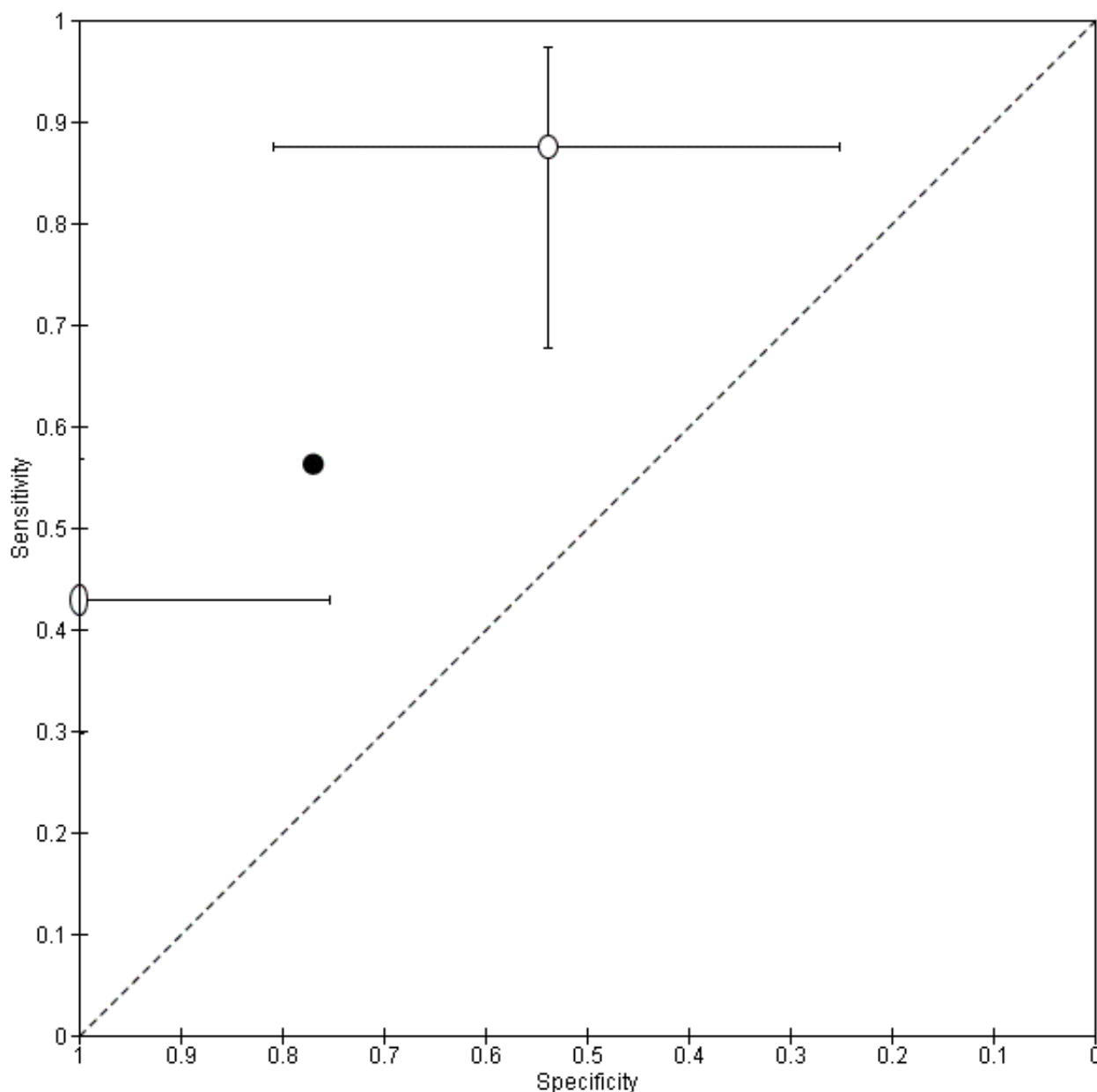


Figure 56. Summary ROC plot of DCBE for detection of rectosigmoid involvement. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. The solid black circle represents the pooled sensitivity and specificity, and the bars correspond to 95% CIs of each individual study.



Indirect comparisons of imaging tests for rectosigmoid endometriosis

TVUS-BP (two studies, 288 participants) (Abrao 2007; Goncalves 2010) demonstrated the highest sensitivity (0.98, 95% CI 0.91 to 1.00 and 0.9 to 1.00; specificity 1.00, 95% CI 0.97 to 1.00 and 0.93 to 1.00) of all TVUS methods. The highest diagnostic estimates of all MRI methods included 3.0T MRI (one study, 41 participants) (sensitivity 1.00, 95% CI 0.75 to 1.00; specificity 0.96, 95% CI 0.82 to 1.00) (Hottat 2009) and MRI 'jelly method' of introducing ultrasonographic gel into both the rectum and the vagina (one study, 260 participants) (sensitivity 0.99, 95% CI 0.96 to 1.00; specificity 0.96, 95% CI 0.90 to 0.99) (Biscaldi 2014). TVUS, TRUS and MRI met the criteria for for

a SpPin triage test and approached the criteria for a SnNout triage test; all demonstrated comparable diagnostic estimates. MDCT-e displayed the best diagnostic performance and met the criteria for both SpPin and SnNout triage tests; however, only three studies examined MDCT-e, and further work is required to confirm these findings. Data for DCBE were scant but largely discouraging.

Direct comparisons of imaging tests for rectosigmoid endometriosis

- 2D-TVUS (sensitivity 0.95, 95% CI 0.87 to 0.99; specificity 0.93, 95% CI 0.87 to 0.97) appeared to be more sensitive and less specific than 3D-TVUS (sensitivity 0.91, 95% CI 0.82 to 0.96;

- specificity 0.97, 95% CI 0.92 to 0.99) for diagnosing rectosigmoid endometriosis in one paired study of 202 participants (Guerriero 2014) (Appendix 8: Figure 57).
- The study that directly compared TVUS, TRUS and MRI (92 participants) (Bazot 2009) revealed that TVUS had higher diagnostic values (sensitivity 0.94, 95% CI 0.85 to 0.98; specificity 1.00, 95% CI 0.88 to 1.00) when compared with MRI (sensitivity 0.87, 95% CI 0.77 to 0.94; specificity 0.93, 95% CI 0.77 to 0.99) and TRUS (sensitivity 0.89, 95% CI 0.78 to 0.95; specificity 0.89, 95% CI 0.78 to 0.95); MRI and TRUS yielded comparable estimates (Appendix 8: Figure 58; Figure 59; Figure 60). This finding was in agreement with two other studies that reported similar types of paired data for detection of rectosigmoid endometriosis (presented below).
 - TVUS (sensitivity 0.96, 95% CI 0.87 to 1.00; specificity 0.90, 95% CI 0.55 to 1.00) was more sensitive and specific than TRUS (sensitivity 0.88, 95% CI 0.76 to 0.96; specificity 0.80, 95% CI 0.44 to 0.97) in a study of 61 participants (Bergamini 2010) (Appendix 8: Figure 61).
 - Further, TVUS (sensitivity 0.98, 95% CI 0.90 to 1.00; specificity 1.00, 95% CI 0.93 to 1.00) was more sensitive and specific than MRI (sensitivity 0.83, 95% CI 0.71 to 0.92; specificity 0.98, 95% CI 0.89 to 1.00) in another direct comparison in 104 participants (Abrao 2007) (Appendix 8: Figure 58).
 - TVUS had higher sensitivity (0.91, 95% CI 0.80 to 0.97) than DCBE (0.43, 95% CI 0.30 to 0.57), although both methods displayed identically high specificity (1.00, 95% CI 0.75 to 1.00) in another head-to-head comparison of 69 participants (Savelli 2011) (Appendix 8: Figure 62).
 - Estimates for TRUS (sensitivity 1.00, 95% CI 0.87 to 1.00; specificity 0.90, 95% CI 0.55 to 1.00) were higher than those for DCBE (sensitivity 0.88, 95% CI 0.68 to 0.97; specificity 0.54, 95% CI 0.25 to 0.81) in a separate direct comparison of 37 participants (Ribeiro 2008a) (Appendix 8: Figure 63).
 - Another paired study (96 participants) (Ferrero 2011) showed that TVUS (RWC-TVS) (sensitivity 0.94, 95% CI 0.83 to 0.99; specificity 0.98, 95% CI 0.89 to 1.00) had lower accuracy estimates than MDCT-e (sensitivity 0.96, 95% CI 0.86 to 0.99; specificity 1.00, 95% CI 0.93 to 1.00) in diagnosing rectosigmoid endometriosis, although both methods demonstrated reasonably high values with overlapping confidence intervals (Appendix 8: Figure 64).
 - MDCT-e (sensitivity 0.99, 95% CI 0.97 to 1.00; specificity 0.99, 95% CI 0.94 to 1.00) and MRI (sensitivity 0.99, 95% CI 0.96 to 1.00; specificity 0.96, 95% CI 0.90 to 0.99) yielded similarly high diagnostic accuracy estimates in one comparative study (260 participants) (Biscaldi 2014) (Appendix 8: Figure 65).
 - 2D-MRI (sensitivity 0.85, 95% CI 0.55 to 0.98; specificity 1.00, 95% CI 0.69 to 1.00) demonstrated similar sensitivity and higher specificity than 3D-MRI (sensitivity 0.85, 95% CI 0.55 to 0.98; specificity 0.90, 95% CI 0.55 to 1.00) in a paired comparative study of 23 participants (Bazot 2013) (Appendix 8: Figure 66).

Figure 57. Forest plot demonstrating the direct comparison between TVUS and 3D-TVUS for rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 2014	73	9	4	116	tg-TVUS	Europe	0.95 [0.87, 0.99]	0.93 [0.87, 0.97]		

TVUS₃ rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 2014	70	4	7	121	3D-TVUS	Europe	0.91 [0.82, 0.96]	0.97 [0.92, 0.99]		

Figure 58. Forest plot demonstrating the direct comparison between TVUS and MRI for rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Abrao 2007	53	0	1	50	TVUS-BP	South America	0.98 [0.90, 1.00]	1.00 [0.93, 1.00]		
Bazot 2009	59	0	4	29	No	Europe	0.94 [0.85, 0.98]	1.00 [0.88, 1.00]		

MRI rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Abrao 2007	45	1	9	49	T1/T2-w + fat-suppressed/Gd	South America	0.83 [0.71, 0.92]	0.98 [0.89, 1.00]		
Bazot 2009	55	2	8	27	T1/T2-w + fat-suppressed/Gd	Europe	0.87 [0.77, 0.94]	0.93 [0.77, 0.99]		

Figure 59. Forest plot demonstrating the direct comparison between TVUS and TRUS for rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	59	0	4	29	No	Europe	0.94 [0.85, 0.98]	1.00 [0.88, 1.00]		
Bergamini 2010	49	1	2	9	RWC-TVUS	Europe	0.96 [0.87, 1.00]	0.90 [0.55, 1.00]		

TRUS rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	56	2	7	27	No	Europe	0.89 [0.78, 0.95]	0.93 [0.77, 0.99]		
Bergamini 2010	45	2	6	8	No	Europe	0.88 [0.76, 0.96]	0.80 [0.44, 0.97]		

Figure 60. Forest plot demonstrating the direct comparison between TRUS and MRI for rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TRUS rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	56	2	7	27	No	Europe	0.89 [0.78, 0.95]	0.93 [0.77, 0.99]		

MRI rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bazot 2009	55	2	8	27	T1/T2-w + fat-suppressed/Gd	Europe	0.87 [0.77, 0.94]	0.93 [0.77, 0.99]		

Figure 61. Forest plot demonstrating the direct comparison between RWC-TVUS and TRUS for rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS, rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
-------	----	----	----	----	-----------------	-------------------	----------------------	----------------------	----------------------	----------------------

TRUS rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
-------	----	----	----	----	-----------------	-------------------	----------------------	----------------------	----------------------	----------------------

Figure 62. Forest plot demonstrating the direct comparison between TVUS and DCBE for rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Savelli 2011	51	0	5	13	No	Europe	0.91 [0.80, 0.97]	1.00 [0.75, 1.00]		

DCBE rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Savelli 2011	24	0	32	13	No	Europe	0.43 [0.30, 0.57]	1.00 [0.75, 1.00]		

Figure 63. Forest plot demonstrating the direct comparison between TRUS and DCBE for rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TRUS rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Ribeiro 2008a	27	1	0	9	No	South America	1.00 [0.87, 1.00]	0.90 [0.55, 1.00]		

DCBE rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Ribeiro 2008a	21	6	3	7	No	South America	0.88 [0.68, 0.97]	0.54 [0.25, 0.81]		

Figure 64. Forest plot demonstrating the direct comparison between RWC-TVUS and MDCT-e for rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS_r rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)

MDCT-e rectosigmoid

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)

Figure 65. Forest plot demonstrating the direct comparison between MDCT-e and MRI for rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

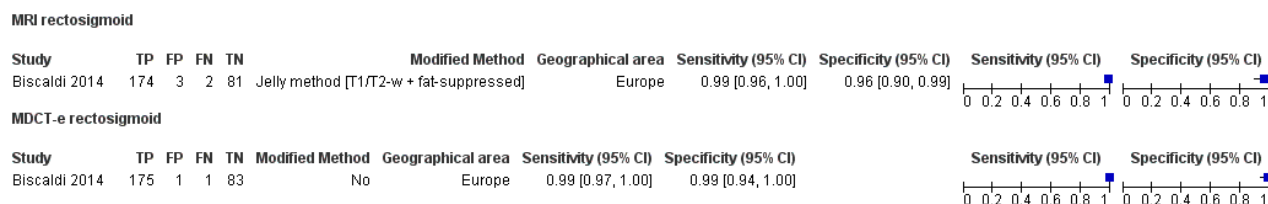
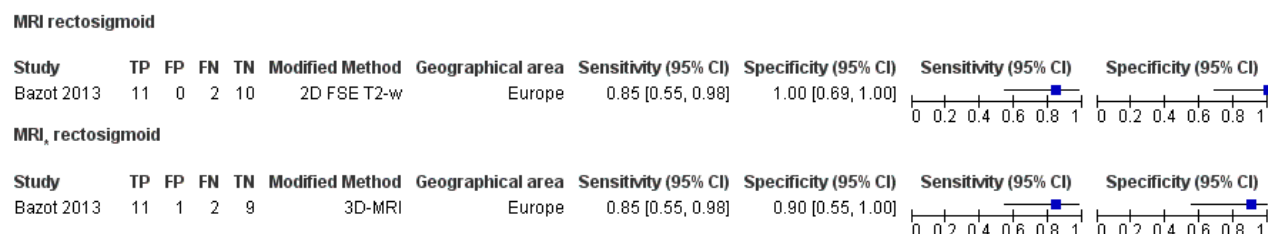


Figure 66. Forest plot demonstrating the direct comparison between 2D-MRI and 3D-MRI for rectosigmoid involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.



Bowel endometriosis (ileum - rectum)

Four studies (six data sets) assessed the accuracy of TVUS, TRUS and MDCT-e in detecting bowel endometriosis from the ileum to the rectum in Europe (n = 3) and Australia (n = 1). For TVUS (three studies, 314 participants), summary sensitivity and specificity were 0.89 (95% CI 0.81 to 0.97) and 0.96 (95% CI 0.91 to 1.00). For TRUS

(one study, 134 participants), sensitivity was 0.96 (95% CI 0.89 to 0.99) and specificity was 1.00 (95% CI 0.94 to 1.00). For MDCT-e (two studies, 194 participants), summary sensitivity and specificity were 0.98 (95% CI 0.92 to 1.00) and 1.00 (95% CI 1.00 to 1.00). Both sensitivity and specificity showed only a small degree of variability; both values were generally high for all tests (Figure 67; Figure 68; Figure 69).

Figure 67. Forest plot of all imaging tests for diagnosis of bowel [ileum - rectum] involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. Studies are ordered according to year of publication for each test. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS bowel [ileum - rectum]

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Piketty 2009	68	2	7	56	No	Europe	0.91 [0.82, 0.96]	0.97 [0.88, 1.00]		
Ferrero 2011	45	1	6	44	RWC-TVS	Europe	0.88 [0.76, 0.96]	0.98 [0.88, 1.00]		
Piessens 2014	22	4	3	56	TVUS-BP	Australia	0.88 [0.69, 0.97]	0.93 [0.84, 0.98]		

TRUS bowel [ileum - rectum]

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Piketty 2009	72	0	3	59	No	Europe	0.96 [0.89, 0.99]	1.00 [0.94, 1.00]		

MDCT-e bowel [ileum - rectum]

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Biscaldi 2007	75	0	1	22	No	Europe	0.99 [0.93, 1.00]	1.00 [0.85, 1.00]		
Ferrero 2011	49	0	2	45	No	Europe	0.96 [0.87, 1.00]	1.00 [0.92, 1.00]		

Figure 68. Summary ROC plot of US methods (TVUS, TRUS) for detection of bowel [ileum - rectum] involvement. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. The solid black circle represents the pooled sensitivity and specificity (for TVUS), and the bars correspond to 95% CIs of each individual study.

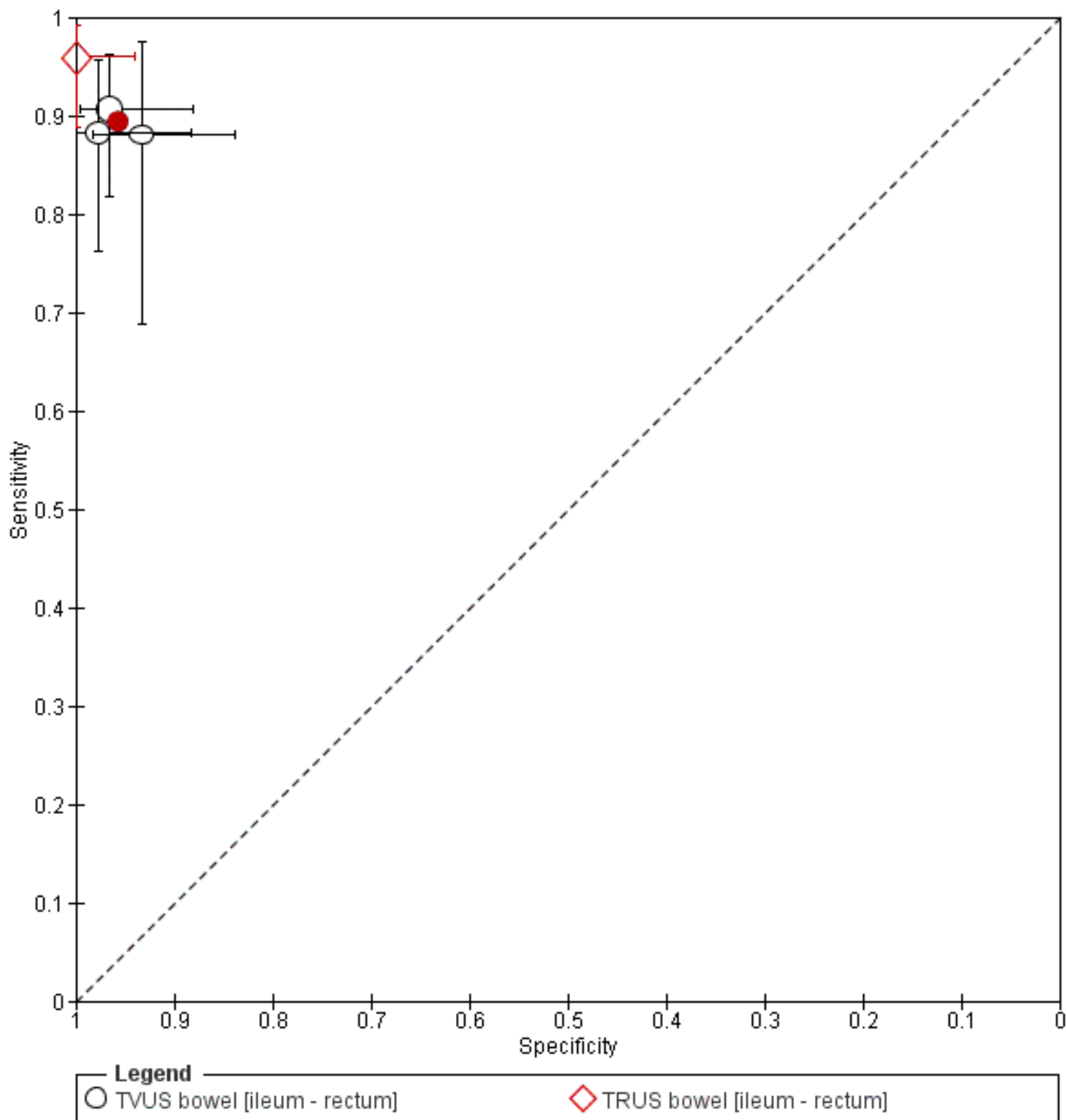
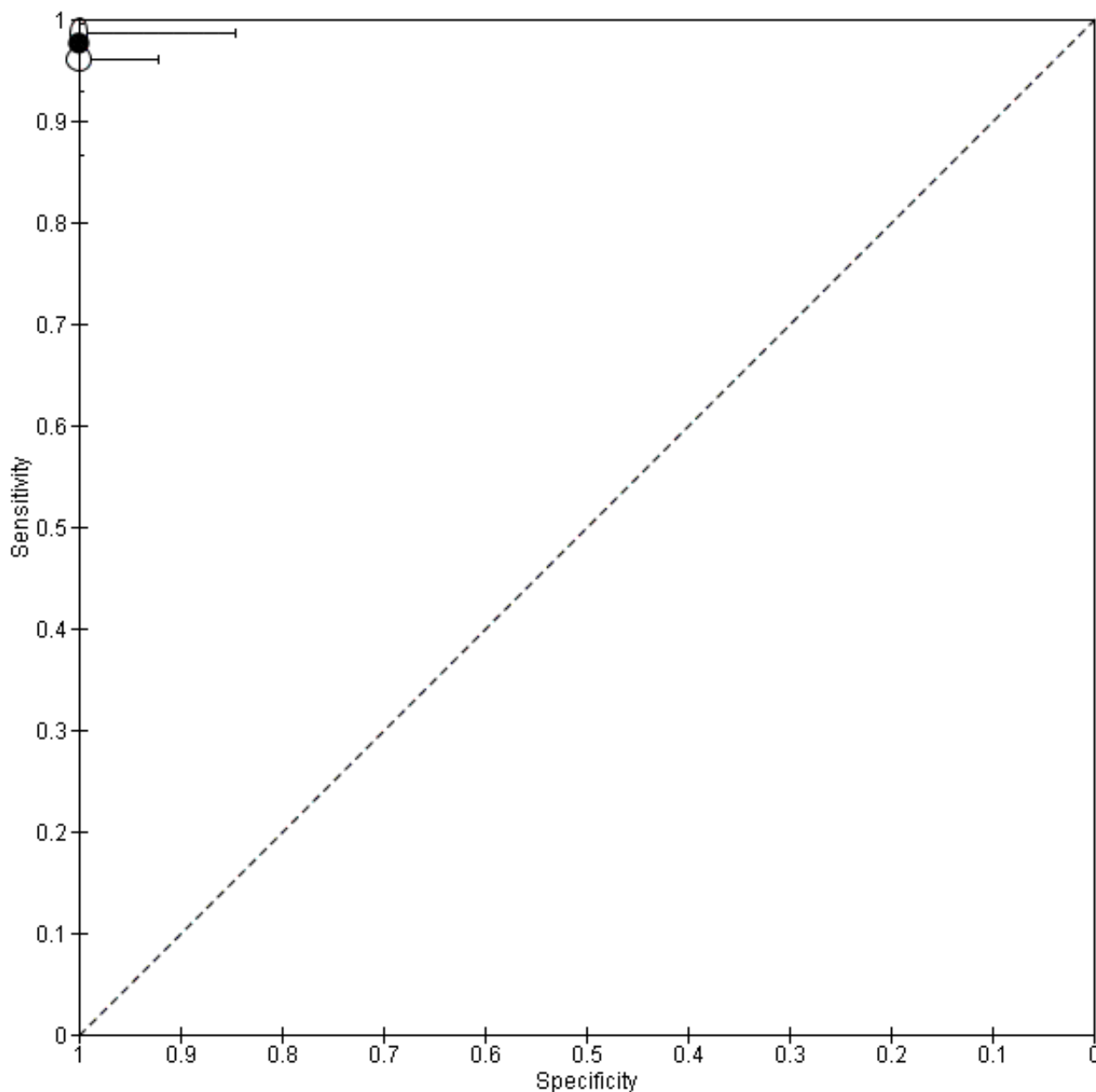


Figure 69. Summary ROC plot of MDCT-e for detection of bowel [ileum - rectum] involvement. Each point represents the pair of sensitivity and specificity from a study. The size of each point is proportional to the study sample size. The solid black circle represents the pooled sensitivity and specificity, and the bars correspond to 95% CIs of each individual study.



Indirect comparisons of imaging tests for bowel endometriosis (ileum - rectum)

The TVUS non-modified technique (one study, 133 participants) (Piketty 2009) showed higher diagnostic estimates than TVUS-BP (one study, 85 participants) and RWC-TVS (one study, 96 participants) with sensitivity of 0.91 (95% CI 0.82 to 0.96) and specificity of 0.97 (95% CI 0.88 to 1.00). Although studies were too few for a meaningful evaluation of the role of imaging tests in diagnosing bowel endometriosis, TVUS, TRUS and MDCT-e met

the criteria for a SpPin triage test, and TRUS and MDCT-e met the criteria for a SnNout triage test for bowel endometriosis.

Direct comparisons of imaging tests for bowel endometriosis (ileum - rectum)

TVUS (sensitivity 0.91, 95% CI 0.82 to 0.96; specificity 0.97, 95% CI 0.88 to 1.00) yielded lower diagnostic accuracy estimates than TRUS (sensitivity 0.96, 95% CI 0.89 to 0.99; specificity 1.00, 95% CI 0.94 to 1.00) in one paired study of 134 participants (Piketty 2009)

(Appendix 8: Figure 70). One study including 96 women (Ferrero 2011) found that MDCT-e (sensitivity 0.96, 95% CI 0.87 to 1.00; specificity 1.00, 95% CI 0.92 to 1.00) had slightly higher estimates

than RWC-TVS (sensitivity 0.88, 95% CI 0.76 to 0.96; specificity 0.98, 95% CI 0.88 to 1.00) for the diagnosis of bowel endometriosis (Appendix 8: Figure 71).

Figure 70. Forest plot demonstrating the direct comparison between TVUS and TRUS for bowel [ileum - rectum] involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS bowel [ileum - rectum]

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Piketty 2009	68	2	7	56	No	Europe	0.91 [0.82, 0.96]	0.97 [0.88, 1.00]		

TRUS bowel [ileum - rectum]

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Piketty 2009	72	0	3	59	No	Europe	0.96 [0.89, 0.99]	1.00 [0.94, 1.00]		

Figure 71. Forest plot demonstrating the direct comparison between RWC-TVS and MDCT-e for bowel [ileum - rectum] involvement. Plot shows study-specific paired estimates of sensitivity and specificity (squares) with 95% CI (black line) and country in which the study was conducted. FN: false negative; FP: false positive; TN: true negative; TP: true positive. Modifications to the conventional technique are presented as 'modified method'.

TVUS bowel [ileum - rectum]

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Ferrero 2011	45	1	6	44	RWC-TVS	Europe	0.88 [0.76, 0.96]	0.98 [0.88, 1.00]		

MDCT-e bowel [ileum - rectum]

Study	TP	FP	FN	TN	Modified Method	Geographical area	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Ferrero 2011	49	0	2	45	No	Europe	0.96 [0.87, 1.00]	1.00 [0.92, 1.00]		

Investigation of heterogeneity and sensitivity analyses

Potential sources of heterogeneity are outlined under [Secondary objectives](#). Although we attempted to assess these sources of heterogeneity, studies evaluating each test were too few to make this a meaningful analysis, except for the meta-analysis with more than 10 studies/data sets (TVUS DIE/posterior DIE, TVUS RVS and TVUS rectosigmoid). For these tests, we found no significant differences in sensitivity or specificity between studies with regards to year of publication, geographical location of the study or application of the modified technique. We were not able to explore the effects of the following potential sources of heterogeneity.

- Age (adolescents vs later reproductive years): information on isolated subgroups not available in any study.
- Clinical presentation (pelvic pain ± infertility vs ovarian mass; symptomatic vs asymptomatic women) or stage of disease (minimal mild, rASRM stage I to II vs moderate to severe, rASRM stage III to IV): information on isolated subgroups not available in any of the studies; all participants symptomatic.
- Histological confirmation versus laparoscopic visualisation without histology: histological test used in conjunction with surgery in most studies.

- Modifications applied to conventional imaging techniques: insufficient number of studies for each method.
- Methodological quality - low versus unclear or high risk: all studies of low methodological quality with high or unclear risk of bias.
- Study design: 'single-gate' versus 'two-gate' studies; all studies except one of single-gate design.

Furthermore, observer variability bias or bias related to interpretation of results cannot be formally assessed in the context of this review.

DISCUSSION

Summary of main results

Data from 4807 women of reproductive age with symptoms of endometriosis who undertook a non-invasive imaging test followed by diagnostic surgery for endometriosis were analysed in 49 articles published from 1993 through 2015. This is the first diagnostic test review to use Cochrane methods and the most comprehensive review to date.

For pelvic endometriosis, no imaging method met the sensitivity criteria for a replacement test or a triage test, albeit TVUS approached the criteria for a SpPin triage test.

For ovarian endometriosis, MRI met the criteria for a replacement test and a SnNout triage test and approached the criteria for a SpPin triage test, but studies were too few to allow conclusions on the role of MRI in detecting ovarian endometriosis. TVUS met the criteria for a SpPin triage test and approached the criteria for a replacement test and a SnNout triage test.

For DIE/posterior DIE, MRI approached the criteria for a replacement test and a SnNout triage test, and TVUS approached the criteria for a SpPin triage test.

Studies were too few for a prudent evaluation of any imaging test for diagnosing anterior DIE.

TVUS, TRUS and MRI reached the criteria for a SpPin triage test and approached the criteria for a SnNout triage test for rectosigmoid endometriosis, which was the most frequently evaluated anatomical site of DIE. TVUS also met SpPin test criteria for other bowel endometriosis (ileum to rectum). MDCT-e displayed the highest diagnostic performance for rectosigmoid and other bowel endometriosis and met the criteria for both SpPin and SnNout triage tests, but studies were too few to provide meaningful results. We found less heterogeneity among estimates for imaging tests in rectosigmoid and bowel endometriosis compared with other anatomical locations, excluding DCBE, which showed heterogeneous and unsatisfactory diagnostic values.

Concerning other anatomical locations, TVUS met the criteria for a SpPin triage test in mapping DIE to USL, RVS, vaginal wall and POD, and MRI could qualify as a SpPin triage test only for POD and vaginal wall endometriosis. TRUS could not be adequately assessed for any of these sites because heterogeneous data were scant.

Data were insufficient for formal comparative analyses between TVUS and MRI methods, although modified ultrasound methods (TVUS-BP and RWC-TVS) and specific MRI modalities (3.0T MRI and MRI jelly method with introduction of ultrasonographic gel into both the rectum and the vagina) showed the highest diagnostic accuracy for evaluated types and anatomical locations of endometriosis.

Studies of poor quality showing considerably heterogeneous results with wide confidence intervals for most evaluated tests suggest caution in interpretation of study results.

Strengths and weaknesses of the review

This review is part of a comprehensive review series of minimally invasive biomarkers for the diagnosis of endometriosis.

Strengths of this review include the following.

- Review authors undertook a very thorough search of the current literature including studies written in languages other than English.
- Two independent review authors extracted data and used a modified QUADAS-2 tool for quality assessments.
- Stringent selection criteria ensured that eligible studies were prospective, included only symptomatic women of reproductive age and performed the index test before providing results of the

reference test, which minimised the risk of bias in interpretation of index test results.

- Most of the included studies (48/49) were of 'single-gate' design, including only clinically relevant populations.
- We approached authors of studies in an attempt to obtain missing information required to assess eligibility and critically appraise studies.

Limitations of this review include the following.

- Few heterogeneous, small studies performed most of the index tests evaluated. This may undermine the reliability of pooled estimates from the meta-analyses and is likely to have contributed to the marked variability in sensitivity and specificity seen for most index tests. Studies varied with respect to participant preparation, operator experience and imaging equipment used, as well as in the definition of the target condition and the diagnostic criteria for imaging tests. Sources of heterogeneity could not be formally explored for most tests because few studies were available for most evaluations. When assessed, geographical location, prevalence of the target condition and assessed risk of bias did not appear to contribute to variation in results.
- All included studies had high/unclear risk of bias; this, together with considerable heterogeneity among studies, contributed to the low quality of evidence presented in this review.
- Reported prevalence of endometriosis in most studies was generally higher than was previously reported for endometriosis (6% to 10% in the general female population and 35% to 50% among symptomatic women for overall endometriosis ([Giudice 2004](#)); 30% for DIE in symptomatic populations) ([Koninckx and Martin 1994](#)). This may reflect a high level of surgical diagnostic expertise but could be due to preselection of more challenging cases at tertiary referral centres and high risk of patient selection bias in most studies. Selection bias appeared to be reduced but not eliminated by consecutive enrolment of participants; however, information on the method of enrolment was missing from most of the included studies.
- Inappropriate assignment to endometriosis and control groups could not be excluded in many studies and is another weakness of the review. Surgical misdiagnosis is a potential cause of bias, as the number and experience of the surgical team, surgical diagnostic criteria and surgical methods were poorly described in most included studies. We now have a standardised technique for performing laparoscopy, and we recommend that future studies should use this standardised method of undertaking laparoscopy ([Becker 2014](#)). Additionally, we did not confine the studies included in this review to those that reported histological confirmation of endometriotic lesions. Although a recent ESHRE guideline stated that evidence is lacking to support laparoscopy without histology to confirm endometriosis ([Dunselman 2014](#)), the clinical significance of histological verification remains debatable. Diagnosis by surgical visualisation only remains a common clinical practice and can be considered reliable when accurate inspection of the abdominal cavity is performed by experienced surgeons. We chose to include the six (15%) studies that reported only surgical visualisation as the reference standard, and we did not wish to lose this potentially valuable information; however, this decision could impact the accuracy of assignment to case and control groups. Moreover, surgeons were commonly aware of

results of the index imaging test preoperatively, which could potentially contribute to bias in interpretation of the reference standard.

- Only five studies addressed interobserver and intraobserver variability for TVUS, reporting that both 2D- and 3D-TVUS were reliable and reproducible techniques. High levels of interobserver concordance were seen between experienced operators ([Holland 2010](#)) and operators with varying degrees of experience ([Guerriero 2007](#); [Pascual 2013](#); [Reid 2013b](#); [Guerriero 2014](#)). For MRI, interobserver agreement varied, with greater intraobserver agreement noted for expert readers and less agreement for junior readers ([Bazot 2013](#)). The diagnostic concordance of observers varied with the location of endometriosis, with high interobserver and intraobserver agreement for ovarian endometrioma, rectosigmoid and RVS disease, and less agreement for identification of uterosacral ligament lesions ([Saba 2010](#); [Bazot 2011b](#); [Saba 2014b](#)).
- Methods for systematic reviews of diagnostic accuracy are emerging, and no criteria for replacement or triage diagnostic tests have been established. We chose criteria that were both realistic and clinically applicable to assist in interpretation of complex results. For a replacement test, we considered the threshold reported by the one and the only systematic review on accuracy of the reference standard (laparoscopy) in detecting endometriosis ([Wykes 2004](#)) to be the most objective. The meta-analysis was published in 2004 and included four eligible studies comprising 433 women. We acknowledge the limitations associated with emphasising a single review, particularly if it does not present the latest and possibly more accurate data that reflect advances in surgical expertise and technology. Several studies on the accuracy of laparoscopy in detecting endometriosis have been published over the past decade; however, their results were not addressed in a systematic way. A further systematic analysis to determine the accuracy of laparoscopy was beyond the scope of this review. Criteria for triage tests utilised the common concepts of SnNout and SpPin in medical statistics, and cut-offs were set at levels that we considered to be clinically relevant (see [Role of index test\(s\)](#)). We encourage the readers to apply independent interpretation of the diagnostic estimates presented while using thresholds that may be more applicable to specific populations and clinical circumstances.

Applicability of findings to the review question

Most studies used QUADAS-2 to rank clinical applicability as high (only one study presented high concern for applicability with regard to patient selection). This reflects inclusion criteria ensuring that prospective symptomatic cohorts of women constituted the participant population, which is highly applicable to the review question and to clinical practice. Most included studies were conducted at specialised centres for endometriosis with a high level of expertise in gynaecological imaging, and index test outcome measures may not be reproducible in all institutions or may not be extrapolated to general practice.

We excluded some potentially relevant well-designed studies as they did not directly address the review question. These included studies that reported the number of endometriotic lesions instead of the number of affected participants as an endpoint. Studies that compared endometriomas versus other ovarian masses did not meet our inclusion criteria for reproductive age or

assessed numbers of cysts rather than numbers of participants. Despite well-defined radiological criteria, endometriomas can be misdiagnosed because of their complex echo texture and multifaceted appearance, and their appearance can be different among premenopausal and postmenopausal women ([Exacoustos 2014](#)). We also excluded rare forms of endometriosis, such as that involving the bladder, ureter or extrapelvic sites (e.g. umbilicus, hernia sacs, abdominal wall, lung, kidney), as studies are informed predominantly by case reports or small case series, and diagnostic laparoscopy is not an applicable reference test for these conditions.

AUTHORS' CONCLUSIONS

Implications for practice

Transvaginal ultrasound (TVUS), the most studied technique, showed only moderate sensitivity, albeit high specificity for pelvic endometriosis and DIE. For these conditions, TVUS did not qualify as a replacement test or a triage test but approached the criteria for a SpPin triage test. In this review, the sensitivity and specificity of TVUS for detecting ovarian endometriosis were high but met the criteria only for a SpPin triage test. In clinical practice, this may mean that the presence of endometriosis (pelvic, ovarian, DIE) on TVUS could establish the diagnosis with high certainty, whereas no radiological evidence of the disease could not confirm that participants are disease-free. This is consistent with international guidelines, which recommend TVUS as first-line investigation in conjunction with history and pelvic examination among women with suspected endometriosis, but do not recommend its use as a replacement test for diagnostic surgery ([ACOG Committee on Gynecology 2010](#); [SOGC 2010](#); [Dunselman 2014](#)). Publications from the past decade suggest that TVUS could accurately detect ovarian endometriosis and could qualify as a replacement test. This theory can be attributed to improved technology and growing experience and should be further validated by use of universal diagnostic criteria and refined radiological protocols.

MRI appeared to be less accurate for peritoneal disease and hence could not qualify as a clinically useful test to replace surgery for overall pelvic endometriosis, but it approached the diagnostic criteria for a replacement test for DIE. Although MRI met the criteria for a replacement test for ovarian endometriosis, evidence is scant and these findings need to be confirmed in larger numbers of studies. In practice, this means that MRI could be utilised in populations for which the risk/benefit ratio of surgery is unclear, such as adolescents, women with significant medical conditions or women with infertility but few pain symptoms of endometriosis. Conservative treatment like the continuous combined oral contraceptive pill or alternative treatments like IVF would be reasonable to consider before surgery. Although guidelines from multiple authorities suggest medical management as first-line treatment for pelvic pain, most women would prefer to receive a definitive diagnosis before commencing potentially long-term therapy. If therapeutic surgery is considered, reliable detection of ovarian endometriomas potentially enables surgeons to assess ovarian reserve and counsel women about fertility preservation before operating on ovarian tissue and risking a reduction in future fertility. Reliably detecting DIE could add weight to a decision to prioritise surgery, and the complexity of surgery and increased risk of complications could be discussed with the woman at the time a decision is needed to undertake surgery.

For most specific anatomical sites of DIE, results of meta-analyses suggest that TVUS could qualify as a SpPin triage test for most anatomical sites, and MRI could be utilised as a SpPin test only for POD, vaginal wall and rectosigmoid endometriosis. Currently MRI is not recommended for routine use in women with endometriosis, but it has been advocated for those with equivocal ultrasound results, for whom rectovaginal or bladder endometriosis is suspected ([ACOG Committee on Gynecology 2010](#)). We did not evaluate bladder endometriosis, but it is interesting to note that MRI did not reach the predetermined diagnostic criteria for USL and RVS endometriosis, and we did not have sufficient data to allow a recommendation on the use of MRI for anterior compartment endometriosis. The clinical utility of a reasonably reliable diagnosis of posterior compartment endometriosis could inform surgeons of the need for a general surgical presence and bowel preparation before the time of surgery. This is particularly important for detecting rectosigmoid endometriosis, as presurgical bowel preparation and surgeries that combine the expertise of gynaecologists and colorectal surgeons (or involve gynaecological surgeons with the expertise to undertake bowel surgery) can be planned preoperatively as rectosigmoid lesions are relatively reliably detected. Rectosigmoid endometriotic lesions were detected with TVUS, TRUS, MRI and MDCT-e with sufficient accuracy (SpPin criteria for TVUS, MRI, TRUS; SpPin and SnNout criteria for MDCT-e). Although studies were too few to allow meaningful evaluation of imaging tests used to detect other bowel endometriosis, small individual studies of TVUS, TRUS and MDCT-e displayed similar performance to that demonstrated for rectosigmoid endometriosis.

We observed that accuracy of the TVUS appeared to be enhanced by bowel preparation (TVUS-BP) and rectal water contrast (RWC-TVS), whereas 3.0T MRI and MRI jelly method with introduction of ultrasonographic gel into both the rectum and the vagina yielded very high diagnostic estimates compared with other MRI modalities. This was consistent for all anatomical sites of DIE, but none of these methods were evaluated for overall pelvic endometriosis. Ultimately, an adequate imaging test is expected to have high accuracy for both diagnosis of endometriosis and presurgical mapping of DIE at specific anatomical locations to simplify the diagnostic algorithm and to reduce the costs of testing. Therefore, further evaluation of modified TVUS methods and specific MRI modalities for overall endometriosis, including peritoneal disease, and for specific anatomical sites is needed.

Data for TRUS were insufficient to permit meaningful recommendations but did not appear to be superior to those for TVUS for any type or site of endometriosis; this brings its clinical utility into question. This observation is particularly important in view of considerable discomfort for women associated with TRUS compared with TVUS.

Although diagnostic potential has been demonstrated for many imaging tests, none of the evaluated tests can be recommended for routine clinical practice, in view of the level of heterogeneity and the wide confidence intervals reported by most studies. Diagnostic estimates of imaging tests for ovarian, rectosigmoid and bowel endometriosis exhibited less heterogeneity compared with tests for other types and locations of endometriosis; this suggests greater reliability, although high/unclear risk of bias in all included studies undermines the reliability of presented results in terms of their clinical utility. We suggest cautious interpretation of presented

data, which in our view cannot be used to confidently inform clinical practice. We encourage further diagnostic research with a focus on potential diagnostic tests identified in this review, in accordance with suggestions presented below for improving the quality of diagnostic research in this field.

We wish to mention that in the absence of well-established criteria for an adequate diagnostic test, the authors of this review determined the diagnostic criteria for replacement and triage tests in a way that we believe will aid interpretation for clinically active readers. However, we encourage readers to apply different criteria according to individual clinical populations and situations.

Implications for research

Currently randomised controlled treatment trials require women with and without endometriosis to have undergone diagnostic surgery for accurate group allocation. For ethical reasons, therapeutic surgery is usually performed at the same time, potentially biasing treatment trial outcomes. Thus our current inability to diagnose and assess the progression of endometriosis in a non-invasive way is a significant limitation in the advancement of clinical research in endometriosis.

Over the past decade, advanced ultrasonographic techniques specifically designed to identify endometriosis, such as the sliding sign, pelvic organ mobility, tenderness-guided ultrasound and use of rectal water contrast and bowel preparation, have been observed to be associated with improvements in the diagnostic accuracy of TVUS for endometriosis. Furthermore, 3.0T MRI and the MRI 'jelly method' appear to have greater diagnostic accuracy than previous older MRI modalities. Studies on these methods are too few to show their value as replacement tests or triage tests for a laparoscopic diagnosis. Additional well-designed diagnostic studies are required to establish the diagnostic test accuracy and clinical utility of these modern imaging methods.

The QUADAS quality assessment of included studies identified several weaknesses in study design that can impede objective evaluation of findings. We recommend that future authors consider (1) including large cohorts after predefining the sample size via a power calculation ([Liu 2005](#)); (2) focusing on a 'single-gate' design that includes only a clinically relevant population ([Rutjes 2005](#)); (3) utilising a diagnostic accuracy study design that adheres to the recommendations of the Standards for Reporting of Diagnostic Accuracy (STARD) initiative ([Bossuyt 2003](#)); (4) incorporating the QUADAS checklist into the study design ([Whiting 2011](#)); (5) formally assessing interobserver and intraobserver variability; (6) establishing universally acceptable diagnostic criteria and radiological protocols; (7) utilising universally acceptable methods of performing laparoscopy ([Becker 2014](#)) as the reference standard test; (8) implementing validation techniques to assess how the results of a statistical analysis will generalise to an independent data set; (9) undertaking direct comparisons of promising tests in conjunction with cost-effectiveness analyses; (10) applying testing to different clinical phenotypes ([Vitonis 2014](#)) rather than to women classified according to rASRM staging; and (11) assessing long-term outcomes and lifetime healthcare costs of women who have participated in diagnostic test accuracy trials of specific diagnostic tests.

Specific opportunities for further research identified by this review include the following.

- Evaluating the ability of TVUS and 3.0T MRI and/or MRI 'jelly method' to diagnose pelvic ovarian endometriosis and DIE/posterior DIE in larger high-quality studies, utilising direct comparisons between methods in conjunction with cost-effectiveness analyses.
- Comparatively evaluating the diagnostic test accuracy of TVUS, TVUS-BP and RWC-TVS in detecting any type of endometriosis.
- Assessing the diagnostic potential of MDCT-e as opposed to other methods in detecting DIE/posterior DIE, rectosigmoid and bowel endometriotic lesions in larger high-quality studies.
- Exploring the value of sequential testing and implementing SnNout and SpPin triage tests for diagnosing endometriosis in conjunction with a cost-effectiveness evaluation of such testing.
- Assessing short- and long-term outcomes and lifetime healthcare costs of women in diagnostic test accuracy trials that have evaluated specific diagnostic imaging tests.

ACKNOWLEDGEMENTS

We would like to thank Associate Professor Petra Macaskill for valuable comments and substantial contributions to development of the statistical methods for this review. Sincere thanks to Professor Ali Akoum and Professor Ian Fraser for intellectual input and assistance with drafting of the protocol. We are grateful to Marian Showell, the Trials Search Co-ordinator of the Cochrane Gynaecology and Fertility Group, for help in designing and conducting the literature search and in locating the full texts of relevant studies. Finally, we thank all contacted study authors who contributed information to this review.

REFERENCES

References to studies included in this review

Abrao 2007 {published data only}

Abrao MS, Goncalves MODC, Dias JA Jr, Podgaec S, Chamie LP, Blasbalg R. Comparison between clinical examination, transvaginal sonography and magnetic resonance imaging for the diagnosis of deep endometriosis. *Human Reproduction* 2007;**22**:3092-7.

Ascher 1995 {published data only}

Ascher SM, Agrawal R, Bis KG, Brown ED, Maximovich A, Markham SM, et al. Endometriosis: appearance and detection with conventional and contrast-enhanced fat-suppressed spin-echo techniques. *Journal of Magnetic Resonance Imaging: JMIR* 1995;**5**(3):251-7.

Bazot 2009 {published data only}

Bazot M, Lafont C, Rouzier R, Roseau G, Thomassin-Naggara I, Darai E. Diagnostic accuracy of physical examination, transvaginal sonography, rectal endoscopic sonography, and magnetic resonance imaging to diagnose deep infiltrating endometriosis. *Fertility and Sterility* 2009;**92**(6):1825-33.

Bazot 2013 {published data only}

Bazot M, Stivalet A, Darai E, Coudray C, Thomassin-Naggara I, Poncelet E. Comparison of 3D and 2D FSE T2-weighted MRI in the diagnosis of deep pelvic endometriosis: preliminary results. *Clinical Radiology* 2013;**68**(1):47-54.

Bergamini 2010 {published data only}

Bergamini V, Ghezzi F, Scarperi S, Raffaelli R, Cromi A, Franchi M. Preoperative assessment of intestinal endometriosis: a comparison of transvaginal sonography with water-contrast in the rectum, transrectal sonography, and barium enema. *Abdominal Imaging* 2010;**35**:732-6.

Biscaldi 2007 {published data only}

Biscaldi E, Ferrero S, Fulcheri E, Ragni N, Remorgida V, Rollandi GA. Multislice CT enteroclysis in the diagnosis of bowel endometriosis. *European Radiology* 2007;**17**:211-9.

Biscaldi 2014 {published data only}

Biscaldi E, Ferrero S, Maggiore ULR, Remorgida V, Venturini PL, Rollandi GA. Multidetector computerized tomography enema versus magnetic resonance enema in the diagnosis of rectosigmoid endometriosis. *European Journal of Radiology* 2014;**83**(2):261-7.

Chamie 2009a {published data only}

Chamie LP, Blasbalg R, Goncalves MOC, Carvalho FM, Abrao MS, de Oliveira IS. Accuracy of magnetic resonance imaging for diagnosis and preoperative assessment of deeply infiltrating endometriosis. *International Journal of Gynecology and Obstetrics* 2009;**106**:198-201.

Dessole 2003 {published data only}

Dessole S, Farina M, Rubattu G, Cosmi E, Ambrosini G, Nardelli GB. Sonovaginography is a new technique for

assessing rectovaginal endometriosis. *Fertility and Sterility* 2003;**79**:1023-7.

Eskenazi 2001 {published data only}

Eskenazi B, Warner M, Bonsignore L, Olive D, Samuels S, Vercellini P. Validation study of nonsurgical diagnosis of endometriosis. *Fertility and Sterility* 2001;**76**:929-35.

Falco 2011 {published data only}

Falco ML, Pareto AE, Serino C, Trezza F, Fusco R, De Luca G, et al. Role of trans-vaginal sonography in deep infiltrating posterior endometriosis [Ruolo dell'ecografia transvaginale nell'endometriosi profonda infiltrante posteriore]. *Giornale Italiano di Ostetricia e Ginecologia* 2011;**33**(4):209-14.

Fastrez 2011 {published data only}

Fastrez M, Nogarede C, Tondeur M, Sirtaine N, Rozenberg S. Evaluation of 18FDG PET-CT in the diagnosis of endometriosis: a prospective study. *Reproductive Sciences* 2011;**18**:540-4.

Fedele 1998 {published data only}

Fedele L, Bianchi S, Portuese A, Borruto F, Dort AM. Transrectal ultrasonography in the assessment of rectovaginal endometriosis. *Obstetrics and Gynecology* 1998;**91**:444-8.

Ferrero 2011 {published data only}

Ferrero S, Biscaldi E, Morotti M, Venturini PL, Remorgida V, Rollandi GA, et al. Multidetector computerized tomography enteroclysis vs. rectal water contrast transvaginal ultrasonography in determining the presence and extent of bowel endometriosis. *Ultrasound in Obstetrics and Gynecology* 2011;**37**(5):603-13.

Ghezzi 2005 {published data only}

Ghezzi F, Raio L, Cromi A, Duwe DG, Beretta P, Buttarelli M, et al. "Kissing ovaries": a sonographic sign of moderate to severe endometriosis. *Fertility and Sterility* 2005;**83**(1):143-7.

Goncalves 2010 {published data only}

Goncalves MODC, Podgaec S, Dias JA, Gonzalez M, Abrao MS. Transvaginal ultrasonography with bowel preparation is able to predict the number of lesions and rectosigmoid layers affected in cases of deep endometriosis, defining surgical strategy. *Human Reproduction* 2010;**25**:665-71.

Grasso 2010 {published data only}

Grasso RF, Di Giacomo V, Sedati P, Sizzi O, Florio G, Faiella E, et al. Diagnosis of deep infiltrating endometriosis: accuracy of magnetic resonance imaging and transvaginal 3D ultrasonography. *Abdominal Imaging* 2010;**35**(6):716-25.

Guerriero 1996a {published data only}

Guerriero S, Ajossa S, Paoletti AM, Mais V, Angiolucci M, Melis GB. Tumor markers and transvaginal ultrasonography in the diagnosis of endometrioma. *Obstetrics and Gynecology* 1996;**88**:403-7.

Guerriero 1996b {published data only}

Guerriero S, Mais V, Ajossa S, Paoletti AM, Angiolucci M, Melis GB. Transvaginal ultrasonography combined with CA-125 plasma levels in the diagnosis of endometrioma. *Fertility and Sterility* 1996;**65**:293-8.

Guerriero 2007 {published data only}

Guerriero S, Ajossa S, Gerada M, D'Aquila M, Piras B, Melis GB. "Tenderness-guided" transvaginal ultrasonography: a new method for the detection of deep endometriosis in patients with chronic pelvic pain. *Fertility and Sterility* 2007;**88**:1293-7.

Guerriero 2008 {published data only}

Guerriero S, Ajossa S, Gerada M, Virgilio B, Angioni S, Melis GB. Diagnostic value of transvaginal 'tenderness-guided' ultrasonography for the prediction of location of deep endometriosis. *Human Reproduction* 2008;**23**:2452-7.

Guerriero 2014 {published data only}

Guerriero S, Saba A, Ajossa S, Peddes C, Angiolucci M, Perniciano M, et al. Three-dimensional ultrasonography in the diagnosis of deep endometriosis. *Human Reproduction* 2014;**29**(6):1189-98. [DOI: [10.1093/humrep/deu054](https://doi.org/10.1093/humrep/deu054)]

Ha 1994 {published data only}

Ha HK, Lim YT, Kim HS, Suh TS, Song HH, Kim SJ. Diagnosis of pelvic endometriosis: fat-suppressed T1-weighted vs conventional MR images. *AJR American Journal of Roentgenology* 1994;**163**:127-31.

Holland 2010 {published data only}

Holland TK, Yazbek J, Cutner A, Saridogan E, Hoo WL, Jurkovic D. Value of transvaginal ultrasound in assessing severity of pelvic endometriosis. *Ultrasound in Obstetrics and Gynecology* 2010;**36**:241-8.

Hottat 2009 {published data only}

Hottat N, Larrousse C, Anaf V, Noel JC, Matos C, Absil J, et al. Endometriosis: contribution of 3.0-T pelvic MR imaging in preoperative assessment - initial results. *Radiology* 2009;**253**:126-34.

Hudelist 2011a {published data only}

Hudelist G, Ballard K, English J, Wright J, Banerjee S, Mastoroudes H, et al. Transvaginal sonography vs. clinical examination in the preoperative diagnosis of deep infiltrating endometriosis. *Ultrasound in Obstetrics and Gynecology* 2011;**37**:480-7.

Hudelist 2013 {published data only}

Hudelist G, Fritzer N, Staettner S, Tammaa A, Tinelli A, Sparic R, et al. Uterine sliding sign: a simple sonographic predictor for presence of deep infiltrating endometriosis of the rectum. *Ultrasound in Obstetrics & Gynecology* 2013;**41**(6):692-5.

Leon 2014 {published data only}

Leon M, Vaccaro H, Alcazar JL, Martinez J, Gutierrez J, Amor F, et al. Extended transvaginal sonography in deep infiltrating endometriosis: use of bowel preparation and an acoustic window with intravaginal gel: preliminary results. *Journal of Ultrasound in Medicine* 2014;**33**(2):315-21.

Manganaro 2012a {published data only}

Manganaro L, Fierro F, Tomei A, Irimia D, Lodise P, Sergi ME, et al. Feasibility of 3.0 T pelvic MR imaging in the evaluation of endometriosis. *European Journal of Radiology* 2012;**81**(6):1381-7.

Manganaro 2012b {published data only}

Manganaro L, Vittori G, Vinci V, Fierro F, Tomei A, Lodise P, et al. Beyond laparoscopy: 3-T magnetic resonance imaging in the evaluation of posterior cul-de-sac obliteration. *Magnetic Resonance Imaging* 2012;**30**(10):1432-8.

Manganaro 2013 {published data only}

Manganaro L, Vinci V, Bernardo S, Storelli P, Fuggetta E, Sollazzo P, et al. The role of 3.0T MRI in the assessment of deep endometriosis located on the uterosacral ligaments. *Journal of Endometriosis* 2013;**5**(1):10-6.

Mangler 2013 {published data only}

Mangler M, Medrano N, Bartley J, Mechsner S, Speiser D, Schneider A, et al. Value of diagnostic procedures in rectovaginal endometriosis. The Australian and New Zealand Journal of Obstetrics and Gynaecology 2013; Vol. 53, issue 4:389-94.

Menada 2008a {published data only}

Menada MV, Remorgida V, Abbamonte LH, Nicoletti A, Ragni N, Ferrero S. Does transvaginal ultrasonography combined with water-contrast in the rectum aid in the diagnosis of rectovaginal endometriosis infiltrating the bowel?. *Human Reproduction* 2008;**23**:1069-75.

Okada 1995 {published data only}

Okada S. Studies on diagnosis of endometriosis by magnetic resonance imaging by means of fat saturation technique. *Acta Obstetrica et Gynaecologica Japonica* 1995;**47**:264-70.

Pascual 2010 {published data only}

Pascual MA, Guerriero S, Hereter L, Barri-Soldevila P, Ajossa S, Graupera B, et al. Diagnosis of endometriosis of the rectovaginal septum using introital three-dimensional ultrasonography. *Fertility and Sterility* 2010;**94**(7):2761-5.

Piessens 2014 {published data only}

Piessens S, Healey M, Maher P, Tsaltas J, Rombauts L. Can anyone screen for deep infiltrating endometriosis with transvaginal ultrasound?. *The Australian and New Zealand Journal of Obstetrics and Gynaecology* 2014;**54**(5):462-8.

Piketty 2009 {published data only}

Piketty M, Chopin N, Dousset B, Millischer-Bellaische AE, Roseau G, Leconte M, et al. Preoperative work-up for patients with deeply infiltrating endometriosis: transvaginal ultrasonography must definitely be the first-line imaging examination. *Human Reproduction* 2009;**24**(3):602-7.

Reid 2013a {published data only}

Reid S, Lu C, Casikar I, Reid G, Abbott J, Cario G, et al. Can we predict pouch of Douglas obliteration in women with suspected endometriosis using a new real-time dynamic transvaginal

ultrasound technique: the “sliding sign”. *Ultrasound in Obstetrics and Gynecology* 2013; Vol. 41, issue 6:685-91.

Reid 2014 {published data only}

Reid S, Lu C, Hardy N, Casikar I, Reid G, Cario G, et al. Office gel sonovaginography for the diagnosis of posterior deep infiltrating endometriosis: a multicenter prospective observational study. *Ultrasound in Obstetrics and Gynecology* 2014;**44**(6):710-8. [DOI: [10.1002/uog.13422](https://doi.org/10.1002/uog.13422)]

Ribeiro 2008a {published data only}

Ribeiro HSAA, Ribeiro PA, Rossini L, Rodrigues FC, Donadio N, Aoki T. Double-contrast barium enema and transrectal endoscopic ultrasonography in the diagnosis of intestinal deeply infiltrating endometriosis. *Journal of Minimally Invasive Gynecology* 2008;**15**:315-20.

Said 2014 {published data only}

Said TH, Azzam AZ. Prediction of endometriosis by transvaginal ultrasound in reproductive-age women with normal ovarian size. *Middle East Fertility Society Journal* 2014;**19**(3):197-207.

Savelli 2011 {published data only}

Savelli L, Manuzzi L, Coe M, Mabrouk M, Di Donato N, Venturoli S, et al. Comparison of transvaginal sonography and double-contrast barium enema for diagnosing deep infiltrating endometriosis of the posterior compartment. *Ultrasound in Obstetrics and Gynecology* 2011;**38**(4):466-71.

Scarella 2013 {published data only}

Scarella AC, Devoto LC, Villarroel CQ, Inzunza NP, Quilodrán FR, Sovino HS. Transvaginal ultrasound for preoperative detection of deep endometriosis in patients with chronic pelvic pain [Ultrasonido transvaginal para la detección preoperatoria de endometriosis profunda en pacientes con dolor pélvico crónico]. *Revista Chilena de Obstetricia y Ginecología* 2013;**78**(2):114-8.

Stabile 2013 {published data only}

Stabile Ianora AA, Moschetta M, Lorusso F, Lattarulo S, Telegrafo M, Rella L, et al. Rectosigmoid endometriosis: comparison between CT water enema and video laparoscopy. *Clinical Radiology* 2013; Vol. 68, issue 9:895-901.

Stratton 2003 {published data only}

Stratton P, Winkel C, Premkumar A, Chow C, Wilson J, Hearn-Stokes R, et al. Diagnostic accuracy of laparoscopy, magnetic resonance imaging, and histopathologic examination for the detection of endometriosis. *Fertility and Sterility* 2003;**79**(5):1078-85.

Sugimura 1993 {published data only}

Sugimura K, Okizuka H, Imaoka I, Kaji Y, Takahashi K, Kitao M, et al. Pelvic endometriosis: detection and diagnosis with chemical shift MR imaging. *Radiology* 1993;**188**(2):435-8.

Takeuchi 2005 {published data only}

Takeuchi H, Kuwatsuru R, Kitade M, Sakurai A, Kikuchi I, Shimanuki H, et al. A novel technique using magnetic resonance imaging jelly for evaluation of rectovaginal endometriosis. *Fertility and Sterility* 2005;**83**(3):442-7.

Thomeer 2014 {published data only}

Thomeer MG, Steensma AB, Van Santbrink EJ, Willemsen FE, Wielopolski PA, Hunink MG, et al. Can magnetic resonance imaging at 3.0-Tesla reliably detect patients with endometriosis? Initial results. *Journal of Obstetrics and Gynaecology Research* 2014;**40**(4):1051-8.

Ubaldi 1998 {published data only}

Ubaldi F, Wisanto A, Camus M, Tournaye H, Clasen K, Devroey P. The role of transvaginal ultrasonography in the detection of pelvic pathologies in the infertility workup. *Human Reproduction* 1998;**13**:330-3.

References to studies excluded from this review

Abrao 2004 {published data only}

Abrao MS, Neme RM, Averbach M, Petta CA, Aldrighi JM. Rectal endoscopic ultrasound with a radial probe in the assessment of rectovaginal endometriosis. *Journal of the American Association of Gynecologic Laparoscopists* 2004;**11**:50-4.

Alcazar 1997 {published data only}

Alcazar JL, Laparte C, Jurado M, Lopez-Garcia G. The role of transvaginal ultrasonography combined with color velocity imaging and pulsed Doppler in the diagnosis of endometrioma. *Fertility and Sterility* 1997;**67**:487-91.

Alcazar 2010 {published data only}

Alcazar JL, Leon M, Galvan R, Guerriero S. Assessment of cyst content using mean gray value for discriminating endometrioma from other unilocular cysts in premenopausal women. *Ultrasound in Obstetrics & Gynecology* 2010;**35**:228-32.

Alcazar 2011 {published data only}

Alcazar JL, Guerriero S, Minguez JA, Ajossa S, Paoletti AM, Ruiz-Zambrana A, et al. Adding cancer antigen 125 screening to gray scale sonography for predicting specific diagnosis of benign adnexal masses in premenopausal women: is it worthwhile?. *Journal of Ultrasound in Medicine* 2011;**30**(10):1381-6.

Anaf 2009 {published data only}

Anaf V, El Nakadi I, De Moor V, Coppens E, Zalcmann M, Noel JC. Anatomic significance of a positive barium enema in deep infiltrating endometriosis of the large bowel. *World Journal of Surgery* 2009;**33**(4):822-7.

Arrive 1989 {published data only}

Arrive L, Hricak H, Martin MC. Pelvic endometriosis: MR imaging. *Radiology* 1989;**171**:687-92.

Ayida 1997 {published data only}

Ayida G, Chamberlain P, Barlow D, Koninckx P, Golding S, Kennedy S. A pilot study assessing the use of hysterosalpingo-contrast sonography and magnetic resonance imaging. *Human Reproduction* 1997;**12**(7):1436-9.

Bahr 2006 {published data only}

Bahr A, De Parades V, Gadonneix P, Etienney I, Salet-Lizee D, Villet R, et al. Endorectal ultrasonography in predicting rectal wall infiltration in patients with deep pelvic endometriosis: a

modern tool for an ancient disease. *Diseases of the Colon and Rectum* 2006;**49**:869-75.

Bazot 2003 {published data only}

Bazot M, Detchev R, Cortez A, Amouyal P, Uzan S, Darai E. Transvaginal sonography and rectal endoscopic sonography for the assessment of pelvic endometriosis: a preliminary comparison. *Human Reproduction* 2003;**18**:1686-92.

Bazot 2004a {published data only}

Bazot M, Darai E, Hourani R, Thomassin I, Cortez A, Uzan S, et al. Deep pelvic endometriosis: MR imaging for diagnosis and prediction of extension of disease. *Radiology* 2004;**232**:379-89.

Bazot 2004b {published data only}

Bazot M, Thomassin I, Hourani R, Cortez A, Darai E. Diagnostic accuracy of transvaginal sonography for deep pelvic endometriosis. *Ultrasound in Obstetrics and Gynecology* 2004;**24**:180-5.

Bazot 2007a {published data only}

Bazot M, Bornier C, Dubernard G, Roseau G, Cortez A, Darai E. Accuracy of magnetic resonance imaging and rectal endoscopic sonography for the prediction of location of deep pelvic endometriosis. *Human Reproduction* 2007;**22**:1457-63.

Bazot 2007b {published data only}

Bazot M, Malzy P, Cortez A, Roseau G, Amouyal P, Darai E. Accuracy of transvaginal sonography and rectal endoscopic sonography in the diagnosis of deep infiltrating endometriosis. *Ultrasound in Obstetrics and Gynecology* 2007;**30**:994-1001.

Bazot 2011a {published data only}

Bazot M, Gasner A, Ballester M, Darai E. Value of thin-section oblique axial T2-weighted magnetic resonance images to assess uterosacral ligament endometriosis. *Human Reproduction* 2011;**26**(2):346-53.

Bazot 2011b {published data only}

Bazot M, Gasner A, Lafont C, Ballester M, Darai E. Deep pelvic endometriosis: limited additional diagnostic value of postcontrast in comparison with conventional MR images. *European Journal of Radiology* 2011;**80**(3):e331-9. [DOI: [10.1016/j.ejrad.2010.12.006](https://doi.org/10.1016/j.ejrad.2010.12.006)]

Bazot 2012 {published data only}

Bazot M, Jarboui L, Ballester M, Touboul C, Thomassin-Naggara I, Darai E. The value of MRI in assessing parametrial involvement in endometriosis. *Human Reproduction* 2012;**27**(8):2352-8.

Bekiesinska-Figatowska 2014 {published data only}

Bekiesinska-Figatowska M. Magnetic resonance imaging as a non-invasive detection tool for extra ovarian endometriosis - own experience [Rezonans magnetyczny jako nieinwazyjne narzędzie detekcji endometriozы pozajajnikowej - doswiadczenie własne]. *Ginekologia Polska* 2014;**85**(9):658-64.

Benaceraff 2015 {published data only}

Benaceraff BR, Groszmann Y, Hornstein MD, Bromley B. Deep infiltrating endometriosis of the bowel wall: the comet sign. *Journal of Ultrasound in Medicine* 2015;**34**(3):537-42.

Boog 1987 {published data only}

Boog G, Penot P, Momber A. Ultrasound as a diagnostic aid in endometriosis. *Contributions to Gynecology and Obstetrics* 1987;**16**:119-24.

Božidar 2010 {published data only}

Njavro B, Njavro L. Comparison of transvaginal ultrasonography and laparoscopy. *Medica Jadertina* 2010;**40**:97-102.

Brazert 2001 {published data only}

Brazert J, Pietryga M, Jasinski P, Szablonski W, Biczysko R. Diagnostic value of transvaginal ultrasound in the detection of ovarian endometriosis. *Ginekologia Polska* 2001;**72**:358-63.

Busard 2010 {published data only}

Busard MPH, Mijatovic V, van Kuijk C, van der Houwen LEE, Bleeker MCG, Cuesta MA, et al. Magnetic resonance imaging in the evaluation of (deep infiltrating) endometriosis: the value of diffusion-weighted imaging. *Journal of Magnetic Resonance Imaging* 2010;**32**(4):1003-9.

Busard 2011 {published data only}

Busard MPH, Pieters van den Bos IC, Mijatovic V, van Kuijk C, Bleeker MCG, van Waesberghe JHTM. Evaluation of MR diffusion-weighted imaging in differentiating endometriosis infiltrating the bowel from colorectal carcinoma. *European Journal of Radiology* 2011;**81**(6):1376-80. [DOI: [10.1016/j.ejrad.2011.03.038](https://doi.org/10.1016/j.ejrad.2011.03.038)]

Busard 2012 {published data only}

Busard MPH, van der Houwen LEE, Bleeker MCG, van den Bos ICP, Cuesta MA, van Kuijk C, et al. Deep infiltrating endometriosis of the bowel: MR imaging as a method to predict muscular invasion. *Abdominal Imaging* 2012;**37**(4):549-57.

Busard 2014 {published data only}

Busard MP, Pieters-van den Bos IC, van Kuijk C, van Waesberghe JHH. Magnetic resonance imaging of deep infiltrating endometriosis: comparison of 2DT2- and 3DT2-weighted TSE sequences. *Journal of Endometriosis* 2014;**6**(1):34-40.

Carbognin 2006 {published data only}

Carbognin G, Girardi V, Pinali L, Raffaelli R, Bergamini V, Pozzi Mucelli R. Assessment of pelvic endometriosis: correlation of US and MRI with laparoscopic findings. *Radiologia Medica* 2006;**111**:687-701.

Cardoso 2009 {published data only}

Cardoso MM, Werner H Jr, Berardo PT, Coutinho AC Jr, Domingues MNA, Gasparetto EL, et al. Evaluation of agreement between transvaginal ultrasonography and magnetic resonance imaging of the pelvis in deep endometriosis with emphasis on intestinal involvement [Avaliação da concordância entre a ultrassonografia transvaginal e a ressonância magnética da pelve na endometriose profunda, com ênfase para

o comprometimento intestinal]. *Radiologica Brasileira* 2009;**42**(2):89-95.

Chamie 2009b {published data only}

Chamie LP. Pelvic endometriosis: correlation among magnetic resonance imaging, laparoscopy and pathological findings. Doctoral dissertation, University of São Paulo [Endometriose pélvica: aspectos à ressonância magnética e correlação com laparoscopia e anatomia patológica. Tese, Universidade de São Paulo]. *Radiologica Brasileira* 2009; Vol. 42, issue 3:158.

Chapron 1998 {published data only}

Chapron C, Dumontier I, Dousset B, Fritel X, Tardif D, Roseau G, et al. Results and role of rectal endoscopic ultrasonography for patients with deep pelvic endometriosis. *Human Reproduction* 1998;**13**(8):2266-70.

Chapron 2004 {published data only}

Chapron C, Vieira M, Chopin N, Balleyguier C, Barakat H, Dumontier I, et al. Accuracy of rectal endoscopic ultrasonography and magnetic resonance imaging in the diagnosis of rectal involvement for patients presenting with deeply infiltrating endometriosis. *Ultrasound in Obstetrics and Gynecology* 2004;**24**(2):175-9.

de Kroon 2004 {published data only}

de Kroon CD, van der Sandt HAGM, van Houwelingen JC, Jansen FW. Sonographic assessment of non-malignant ovarian cysts: does sonohistology exist?. *Human Reproduction* 2004;**19**:2138-43.

Delpy 2005 {published data only}

Delpy R, Barthet M, Gasmi M, Berdah S, Shojai R, Desjeux A, et al. Value of endorectal ultrasonography for diagnosing rectovaginal septal endometriosis infiltrating the rectum. *Endoscopy* 2005;**37**:357-61.

Demidov 1991 {published data only}

Demidov VN, Strukov AV, Gus AI. Transvaginal echography in the diagnosis of endometrioid cysts of the ovaries. *Vestnik Rentgenologii i Radiologii* 1991;**1**:48-51.

De Souza 1995 {published data only}

De Souza NM, Brosens JJ, Schwieso JE, Paraschos T, Winston RML. The potential value of magnetic resonance imaging in infertility. *Clinical Radiology* 1995;**50**:75-9.

Di Paola 2015 {published data only}

Di Paola V, Manfredi R, Castelli F, Negrelli R, Mehrabi S, Pozzi Mucelli R. Detection and localization of deep endometriosis by means of MRI and correlation with the ENZIAN score. *European Journal of Radiology* 2015;**8**(84):568-74. [DOI: [10.1016/j.ejrad.2014.12.017](https://doi.org/10.1016/j.ejrad.2014.12.017)]

Dogan 1996 {published data only}

Dogan MM, Ugur M, Soysal SK, Soysal ME, Ekici E, Gokmen O. Transvaginal sonographic diagnosis of ovarian endometrioma. *International Journal of Gynecology and Obstetrics* 1996;**52**:145-9.

Drobne 2014 {published data only}

Drobne D, Ribic-Pucelj M, Stepec S, Tosovic Z, Gruden A, Mervic M, et al. Rectal endoscopic ultrasound for the diagnostics of bowel endometriosis [Pomen endoskopskega ultrazvoka v diagnostiki endometrioze crevesa Slovene]. *Zdravniški Vestnik* 2014;**83**(12):857-64.

Dumontier 2000 {published data only}

Dumontier I, Roseau G, Vincent B, Chapron C, Dousset B, Chaussade S, et al. Comparison of endoscopic ultrasound and magnetic resonance imaging in pelvic endometriosis [French]. *Gastroenterologie Clinique et Biologique* 2000;**24**(12):1197-204.

Egekvist 2012 {published data only}

Egekvist AG, Forman A, Seyer-Hansen M. Transvaginal ultrasonography of rectosigmoid endometriosis: interobserver variation of lesion size. *Acta Obstetrica et Gynecologica Scandinavica* 2012;**91**(2):264-8.

Exacoustos 2013 {published data only}

Exacoustos C, Luciano D, Corbett B, De Felice G, Di Felicianantonio M, Luciano A, et al. The uterine junctional zone: a 3-dimensional ultrasound study of patients with endometriosis. *American Journal of Obstetrics and Gynecology* 2013;**209**(3):248.e1-7.

Exacoustos 2014 {published data only}

Exacoustos C, Malzoni M, Di Giovanni A, Lazzeri L, Tosti C, Petraglia F, et al. Ultrasound mapping system for the surgical management of deep infiltrating endometriosis. *Fertility and Sterility* 2014;**102**(1):143-50.e2.

Faccioli 2008 {published data only}

Faccioli N, Manfredi R, Mainardi P, Chiara ED, Spoto E, Minelli L, et al. Barium enema evaluation of colonic involvement in endometriosis. *American Journal of Roentgenology* 2008;**190**(4):1050-4.

Faccioli 2010 {published data only}

Faccioli N, Foti G, Manfredi R, Mainardi P, Spoto E, Ruffo G, et al. Evaluation of colonic involvement in endometriosis: double-contrast barium enema vs. magnetic resonance imaging. *Abdominal Imaging* 2010;**35**:414-21.

Falco 1995 {published data only}

Falco G, De Frenza N, Ventrella C, Traversa M, Chiuri E, Ventrella V. Magnetic resonance imaging in the diagnosis of endometriosis [Italian]. *Giornale Italiano di Ostetricia e Ginecologia* 1995;**17**(1):35-9.

Fiaschetti 2012 {published data only}

Fiaschetti V, Crusco S, Meschini A, Cama V, Di Vito L, Marziali M, et al. Deeply infiltrating endometriosis: evaluation of retro-cervical space on MRI after vaginal opacification. *European Journal of Radiology* 2012;**81**(11):3638-45.

Fratelli 2013 {published data only}

Fratelli N, Scioscia M, Bassi E, Musola M, Minelli L, Trivella G. Transvaginal sonography for preoperative assessment of deep endometriosis. *Journal of Clinical Ultrasound* 2013;**41**(2):69-75.

Friedman 1985 {published data only}

Friedman H, Vogelzang RL, Mendelson EB. Endometriosis detection by US with laparoscopic correlation. *Radiology* 1985;**157**:217-20.

Gauche Cazalis 2012 {published data only}

Cazalis G, Koskas M, Martin B, Palazzo L, Madelenat P, Yazbeck C. Preoperative imaging of deeply infiltrating endometriosis: transvaginal sonography, rectal endoscopic sonography and magnetic resonance imaging [Imagerie préopératoire dans l'endométriose profonde : échographie pelvienne, échographie rectale et IRM]. *Gynécologie, Obstétrique & Fertilité* 2012;**40**(11):634-41.

Gordon 1982 {published data only}

Gordon RL, Evers K, Kressel HY, Laufer I, Herlinger H, Thompson JJ. Double-contrast enema in pelvic endometriosis. *AJR. American Journal of Roentgenology* 1982;**138**:549-52.

Griffiths 2008 {published data only}

Griffiths A, Koutsouridou R, Vaughan S, Penketh R, Roberts SA, Torkington J. Transrectal ultrasound and the diagnosis of rectovaginal endometriosis: a prospective observational study. *Acta Obstetrica et Gynecologica Scandinavica* 2008;**87**:445-8.

Guerriero 1995 {published data only}

Guerriero S, Mais V, Ajossa S, Paoletti AM, Angiolucci M, Labate F, et al. The role of endovaginal ultrasound in differentiating endometriomas from other ovarian cysts. *Clinical and Experimental Obstetrics and Gynecology* 1995;**22**(1):20-2.

Guerriero 1997 {published data only}

Guerriero S, Mallarini G, Ajossa S, Risalvato A, Satta R, Mais V, et al. Transvaginal ultrasound and computed tomography combined with clinical parameters and CA-125 determinations in the differential diagnosis of persistent ovarian cysts in premenopausal women. *Ultrasound in Obstetrics and Gynecology* 1997;**9**:339-43.

Guerriero 1998 {published data only}

Guerriero S, Ajossa S, Mais V, Risalvato A, Lai MP, Melis GB. The diagnosis of endometriomas using colour Doppler energy imaging. *Human Reproduction* 1998;**13**:1691-5.

Guerriero 2009 {published data only}

Guerriero S, Alcazar JL, Pascual MA, Ajossa S, Gerada M, Bargellini R, et al. Diagnosis of the most frequent benign ovarian cysts: is ultrasonography accurate and reproducible?. *Journal of Women's Health* 2009;**18**:519-27.

Guerriero 2010 {published data only}

Guerriero S, Ajossa S, Garau N, Alcazar JL, Mais V, Melis GB. Diagnosis of pelvic adhesions in patients with endometrioma: the role of transvaginal ultrasonography. *Fertility and Sterility* 2010;**94**(2):742-6.

Hauth 2004 {published data only}

Hauth EAM, Antoch C, Ruehm SG, Boing C, Kimmig R, Forsting M. Value of pelvic MRI in the preoperative diagnosis of endometriosis. *RoFo Fortschritte auf dem Gebiet*

der Röntgenstrahlen und der Bildgebenden Verfahren 2004;**176**(9):1265-70.

Hensen 2009 {published data only}

Hensen JHJ, Puylaer JBCM. Endometriosis of the posterior cul-de-sac: clinical presentation and findings at transvaginal ultrasound. *American Journal of Roentgenology* 2009;**192**:1618-24.

Holland 2013a {published data only}

Holland TK, Cutner A, Saridogan E, Mavrellos D, Pateman K, Jurkovic D. Ultrasound mapping of pelvic endometriosis: does the location and number of lesions affect the diagnostic accuracy? A multicentre diagnostic accuracy study. *BMC Womens Health* 2013;**13**:43.

Holland 2013b {published data only}

Holland TK, Hoo WL, Mavrellos D, Saridogan E, Cutner A, Jurkovic D. Reproducibility of assessment of severity of pelvic endometriosis using transvaginal ultrasound. *Ultrasound in Obstetrics & Gynecology* 2013;**41**:210-5.

Hudelist 2009a {published data only}

Hudelist G, Oberwinkler KH, Singer CF, Tuttlies F, Rauter G, Ritter O, et al. Combination of transvaginal sonography and clinical examination for preoperative diagnosis of pelvic endometriosis. *Human Reproduction* 2009;**24**:1018-24.

Hudelist 2009b {published data only}

Hudelist G, Tuttlies F, Rauter G, Pucher S, Keckstein J. Can transvaginal sonography predict infiltration depth in patients with deep infiltrating endometriosis of the rectum?. *Human Reproduction* 2009;**24**:1012-7.

Iosca 2013 {published data only}

Iosca S, Lumia D, Bracchi E, Duka E, De Bon M, Lekaj M, et al. Multislice computed tomography with colon water distension (MSCT-c) in the study of intestinal and ureteral endometriosis. *Clinical Imaging* 2013;**37**(6):1061-8.

Jain 1993 {published data only}

Jain KA, Friedman DL, Pettinger TW, Alagappan R, Jeffrey RB Jr, Sommer FG. Adnexal masses: comparison of specificity of endovaginal US and pelvic MR imaging. *Radiology* 1993;**186**:697-704.

Jarlot 2008 {published data only}

Jarlot C, Anglade E, Paillocher N, Moreau D, Catala L, Aube C. MR imaging features of deep pelvic endometriosis: correlation with laparoscopy. *Journal de Radiologie* 2008;**89**:1745-54.

Jeong 2013 {published data only}

Jeong SY, Chung DJ, Myung YD, Lim YT, Hahn ST, Lee JM. The usefulness of computed tomographic colonography for evaluation of deep infiltrating endometriosis: comparison with magnetic resonance imaging. *Journal of Computer Assisted Tomography* 2013;**37**(5):809-14.

Jermy 2001 {published data only}

Jermy K, Luise C, Bourne T. The characterization of common ovarian cysts in premenopausal women. *Ultrasound in Obstetrics and Gynecology* 2001;**17**:140-4.

Johnson 1994 {published data only}

Johnson W, Ott D, Chen M, Favez J, Gelfand D. Efficacy of hysterosalpingography in evaluating endometriosis. *Abdominal Imaging* 1994;**19**:278-80.

Jung 2010 {published data only}

Jung SI, Kim YJ, Jeon HJ, Jeong KA. Deep infiltrating endometriosis: CT imaging evaluation. *Journal of Computer Assisted Tomography* 2010;**34**(3):338-42.

Khan 2013 {published data only}

Khan AA, Bashir N, Akram R, Shami N, Anwar S, Asif S, et al. Transvaginal sonographic diagnosis of ovarian endometrioma. *Pakistan Journal of Medical and Health Sciences* 2013;**7**(1):22-6.

Kikuchi 2009 {published data only}

Kikuchi I, Takeuchi H, Kuwatsuru R, Kitade M, Kumakiri J, Kuroda K, et al. Diagnosis of complete cul-de-sac obliteration (CCDSO) by the MRI jelly method. *Journal of Magnetic Resonance Imaging* 2009;**29**(2):365-70.

Kikuchi 2014 {published data only}

Kikuchi I, Kuwatsuru R, Yamazaki K, Kumakiri J, Aoki Y, Takeda S. Evaluation of the usefulness of the MRI jelly method for diagnosing complete cul-de-sac obliteration. *BioMed Research International* 2014;**Epub 2014/05/09**:1-7.

Kinkel 1999 {published data only}

Kinkel K, Chapron C, Balleyguier C, Fritel X, Dubuisson JB, Moreau JF. Magnetic resonance imaging characteristics of deep endometriosis. *Human Reproduction* 1999;**14**:1080-6.

Kreuzberg 2004 {published data only}

Kreuzberg B, Kastner J, Novotny Z, Ulcova-Gallova Z, Opatrny V, Mukensnabl P. The contribution of magnetic resonance examination in the diagnosis of endometriosis. *Ceska Radiologie* 2004;**58**(2):79-85.

Kruger 2013 {published data only}

Krüger K, Behrendt K, Niedobitek-Kreuter G, Koltermann K, Ebert AD. Location-dependent value of pelvic MRI in the preoperative diagnosis of endometriosis. *European Journal of Obstetrics & Gynecology and Reproductive Biology* 2013;**169**:93-9.

Kurjak 1994 {published data only}

Kurjak A, Kupesic S. Scoring system for prediction of ovarian endometriosis based on transvaginal color and pulsed Doppler sonography. *Fertility and Sterility* 1994;**62**:81-8.

Li 2012 {published data only}

Li YP, Wang N, Zhang L, Zhu HM, Wang LS, Shi RY. Value of linear endoscopic ultrasonography in the diagnosis of rectal endometriosis. *World Chinese Journal of Digestology* 2012;**20**(14):1252-5.

Li 2014 {published data only}

Li Y, Song QW, Sun MY, Wang HQ, Wang S, Wei Q, et al. Use of enhanced T2 star-weighted angiography (ESWAN) and R2* values to distinguish ovarian cysts due to endometriosis from other causes. *Abdominal Imaging* 2014;**Epub 2014/12/17**:1-9. [10.1095/biolreprod.114.124891]

Macario 2012 {published data only}

Macario S, Chassang M, Novellas S, Baudin G, Delotte J, Toullalan O, et al. The value of pelvic MRI in the diagnosis of posterior cul-de-sac obliteration in cases of deep pelvic endometriosis. *American Journal of Roentgenology* 2012;**199**(6):1410-5.

Mais 1993 {published data only}

Mais V, Guerriero S, Ajossa S, Angiolucci M, Paoletti A M, Melis GB. The efficiency of transvaginal ultrasonography in the diagnosis of endometrioma. *Fertility and Sterility* 1993;**60**:776-80.

Mathlouthi 2011 {published data only}

Mathlouthi N, Ben Ayed B, Dhouib M, Chaabene K, Trabelsi K, Amouri H, et al. [Prospective study of the correlation of ultrasonography and CA125 in the management of ovarian cysts: a study of 77 cases]. *La Tunisie Medicale* 2011;**89**:686-92.

Menada 2008b {published data only}

Menada MV, Remorgida V, Abbamonte LH, Fulcheri E, Ragni N, Ferrero S. Transvaginal ultrasonography combined with water-contrast in the rectum in the diagnosis of rectovaginal endometriosis infiltrating the bowel. *Fertility and Sterility* 2008;**89**:699-700.

Mezzi 2011 {published data only}

Mezzi G, Ferrari S, Arcidiacono PG, Puppo FD, Candiani M, Testoni PA. Endoscopic rectal ultrasound and elastosonography are useful in flow chart for the diagnosis of deep pelvic endometriosis with rectal involvement. *Journal of Obstetrics and Gynaecology Research* 2011;**37**:586-90.

Millischer 2014 {published data only}

Millischer AE, Salomon L, Santulli P, Borghese B, Dousset B, Chapron C. Real-time virtual sonography using MRI-US fusion imaging for the evaluation of deep infiltrating endometriosis: feasibility and preliminary results. *Ultrasound in Obstetrics and Gynecology* 2014;**Epub date: 2014/11/02**:1-23.

Minaif 2008 {published data only}

Minaif K, Shigueoka DC, Minami CCS, Sales DM, Ruano JMC, Noguti AS, et al. Pelvic endometriosis: a comparison between low-field (0.2 T) and high-field (1.5 T) magnetic resonance imaging [Endometriose pélvica: comparação entre imagens por ressonância magnética de baixo campo (0.2 T) e alto campo (1.5 T)]. *Radiologia Brasileira* 2008;**41**(6):367-72.

Nezhat 1994 {published data only}

Nezhat C, Santolaya J, Nezhat FR. Comparison of transvaginal sonography and bimanual pelvic examination in patients with laparoscopically confirmed endometriosis. *The Journal of the American Association of Gynecologic Laparoscopists* 1994;**1**(2):127-30.

Njavro 2003 {published data only}

Njavro B, Hodzic D. Comparison of transvaginal ultrasonography and minilaparoscopy in assessment of benign adnexal masses. *Gynaecologia et Perinatologia* 2003;**12**:122-7.

Ohba 1996 {published data only}

Ohba T, Mizutani H, Maeda T, Matsuura K, Okamura H. Evaluation of endometriosis in uterosacral ligaments by transrectal ultrasonography. *Human Reproduction* 1996;**11**:2014-7.

Okaro 2006 {published data only}

Okaro E, Condous G, Khalid A, Timmerman D, Ameye L, Huffel SV, et al. The use of ultrasound-based 'soft markers' for the prediction of pelvic pathology in women with chronic pelvic pain - Can we reduce the need for laparoscopy?. *BJOG: An International Journal of Obstetrics and Gynaecology* 2006;**113**(3):251-6.

Onbas 2007 {published data only}

Onbas O, Kantarci M, Alper F, Kumtepe Y, Durur I, Ingec M, et al. Nodular endometriosis: dynamic MR imaging. *Abdominal Imaging* 2007;**32**(4):451-6.

Outwater 1993 {published data only}

Outwater E, Schiebler ML, Owen RS, Schnall MD. Characterization of hemorrhagic adnexal lesions with MR imaging: blinded reader study. *Radiology* 1993;**186**:489-94.

Pascual 2000 {published data only}

Pascual MA, Tresserra F, Lopez-Marin L, Ubeda A, Grases PJ, Dexeus S. Role of color doppler ultrasonography in the diagnosis of endometriotic cyst. *Journal of Ultrasound in Medicine* 2000;**19**(10):695-9.

Pascual 2013 {published data only}

Pascual MA, Guerriero S, Hereter L, Barri-Soldevila P, Ajossa S, Gaupera B, et al. Three-dimensional sonography for diagnosis of rectovaginal septum endometriosis interobserver agreement. *Journal of Ultrasound in Medicine* 2013;**32**(6):931-5.

Patel 1999 {published data only}

Patel MD, Feldstein VA, Chen DC, Lipson SD, Filly RA. Endometriomas: diagnostic performance of US. *Radiology* 1999;**210**(3):739-45.

Pereira 2009 {published data only}

Pereira RMA, Zanatta A, de Mello Bianchi PH, Chamie LP, Goncalves MOC, Serafini PC. Transvaginal ultrasound after bowel preparation to assist surgical planning for bowel endometriosis resection. *International Journal of Gynecology and Obstetrics* 2009;**104**(2):161.

Philip 2015 {published data only}

Philip CA, Bisch C, Coulon A, de Saint-Hilaire P, Rudigoz RC, Dubernard G. Correlation between three-dimensional rectosonography and magnetic resonance imaging in the diagnosis of rectosigmoid endometriosis: a preliminary study on the first fifty cases. *European Journal of Obstetrics and Gynecology and Reproductive Biology* 2015;**187**:35-40.

Pishvaian 2006 {published data only}

Pishvaian AC, Ahlawat SK, Garvin D, Haddad NG. Role of EUS and EUS-guided FNA in the diagnosis of symptomatic rectosigmoid endometriosis. *Gastrointestinal Endoscopy* 2006;**63**(2):331-5.

Preutthipan 1995 {published data only}

Preutthipan S, Hesla JS. A comparative study between pelvic ultrasonography and laparoscopy in the detection of pelvic pathology in the initial workup of subfertile women. *Journal of the Medical Association of Thailand = Chotmaihet thangphaet* 1995;**78**:596-9.

Reid 2013b {published data only}

Reid S, Lu C, Casikar I, Mein B, Magotti R, Ludlow J, et al. The prediction of pouch of Douglas obliteration using offline analysis of the transvaginal ultrasound 'sliding sign' technique: inter- and intra-observer reproducibility. *Human Reproduction* 2013;**28**(5):1237-46.

Ribeiro 2008b {published data only}

Ribeiro HSAA, Ribeiro PAG, Rodrigues FC, Donadio N, Auge APF, Aoki T. Double-contrast barium enema in the diagnosis of intestinal deeply infiltrating endometriosis. *Revista Brasileira de Ginecologia e Obstetricia* 2008;**30**:400-5.

Roman 2008 {published data only}

Roman H, Kouteich K, Gromez A, Hochain P, Resch B, Marpeau L. Endorectal ultrasound accuracy in the diagnosis of rectal endometriosis infiltration depth. *Fertility and Sterility* 2008;**90**(4):1008-13.

Roseau 2000 {published data only}

Roseau G, Dumontier I, Palazzo L, Chapron C, Dousset B, Chaussade S, et al. Rectosigmoid endometriosis: endoscopic ultrasound features and clinical implications. *Endoscopy* 2000;**32**(7):525-30.

Rossi 2014 {published data only}

Rossi L, Palazzo L, Yazbeck C, Walker F, Chis C, Luton D, et al. Can rectal endoscopic sonography be used to predict infiltration depth in patients with deep infiltrating endometriosis of the rectum?. *Ultrasound in Obstetrics and Gynecology* 2014;**43**(3):322-7.

Rousset 2014 {published data only}

Rousset P, Peyron N, Charlot M, Chateau F, Golfier F, Raudrant D, et al. Bowel endometriosis: preoperative diagnostic accuracy of 3.0-T MR enterography - initial results. *Radiology* 2014;**273**(1):117-24.

Roy 2009 {published data only}

Roy C, Balzan C, Thoma V, Sauer B, Wattiez A, Leroy J. Efficiency of MR imaging to orientate surgical treatment of posterior deep pelvic endometriosis. *Abdominal Imaging* 2009;**34**(2):251-9.

Saba 2010 {published data only}

Saba L, Guerriero S, Sulcis R, Ajossa S, Melis G, Mallarini G. Agreement and reproducibility in identification of endometriosis using magnetic resonance imaging. *Acta Radiologica* 2010;**51**(5):573-80.

Saba 2011 {published data only}

Saba L, Guerriero S, Sulcis R, Pilloni M, Ajossa S, Melis G, et al. Learning curve in the detection of ovarian and deep endometriosis by using magnetic resonance: comparison with surgical results. *European Journal of Radiology* 2011;**79**(2):237-44.

Saba 2012 {published data only}

Saba L, Guerriero S, Sulcis R, Pilloni M, Ajossa S, Melis G, et al. MRI and "tenderness guided" transvaginal ultrasonography in the diagnosis of recto-sigmoid endometriosis. *Journal of Magnetic Resonance Imaging* 2012; Vol. 35, issue 2:352-60.

Saba 2014b {published data only}

Saba L, Sulcis R, Melis GB, Ibba G, Alcazar JL, Piga M, et al. Diagnostic confidence analysis in the magnetic resonance imaging of ovarian and deep endometriosis: comparison with surgical results. *European Radiology* 2014;**24**(2):335-43.

Saccardi 2012 {published data only}

Saccardi C, Cosmi E, Borghero A, Tregnaghi A, Dessole S, Litta P. Comparison between transvaginal ultrasound, sonovaginography and magnetic resonance imaging in the diagnosis of posterior deep infiltrating endometriosis. *Ultrasound in Obstetrics and Gynecology* 2012; Vol. 40, issue 4:464-9.

Scardapane 2011 {published data only}

Scardapane A, Bettocchi S, Lorusso F, Stabile Ianora AA, Vimercati A, Ceci O, et al. Diagnosis of colorectal endometriosis: contribution of contrast enhanced MR-colonography. *European Radiology* 2011;**21**(7):1553-63.

Scardapane 2013 {published data only}

Scardapane A, Lorusso F, Bettocchi S, Moschetta M, Fiume M, Vimercati A, et al. Deep pelvic endometriosis: accuracy of pelvic MRI completed by MR colonography. *La Radiologia Medica* 2013;**118**(2):323-38.

Scardapane 2014 {published data only}

Scardapane A, Lorusso F, Scioscia M, Ferrante A, Stabile Ianora AA, Angelelli G. Standard high-resolution pelvic MRI vs. low-resolution pelvic MRI in the evaluation of deep infiltrating endometriosis. *European Radiology* 2014;**24**(10):2590-6.

Schroder 1997 {published data only}

Schroder J, Lohnert M, Doniec JM, Dohrmann P. Endoluminal ultrasound diagnosis and operative management of rectal endometriosis. *Diseases of the Colon and Rectum* 1997;**40**(5):614-7.

Setubal 2011 {published data only}

Setubal A, Maia S, Lowenthal C, Sidiropoulou Z. FDG-PET value in deep endometriosis. *Gynecological Surgery* 2011;**8**(3):305-9.

Sherif 2015 {published data only}

Sherif MF, Badawy ME, Elkholy DGEY. Accuracy of magnetic resonance imaging in diagnosis of deeply infiltrating endometriosis. *Egyptian Journal of Radiology and Nuclear Medicine* 2015;**46**(1):159-65.

Sokalska 2009 {published data only}

Sokalska A, Timmerman D, Testa AC, Van Holsbeke C, Lissoni AA, Leone FP, et al. Diagnostic accuracy of transvaginal ultrasound examination for assigning a specific diagnosis to adnexal masses. *Ultrasound in Obstetrics and Gynecology* 2009;**34**(4):462-70.

Stegmann 2009 {published data only}

Stegmann BJ, Funk MJ, Sinaii N, Hartmann KE, Segars J, Nieman LK, et al. A logistic model for the prediction of endometriosis. *Fertility and Sterility* 2009;**91**(1):51-5.

Takahashi 1994 {published data only}

Takahashi K, Okada S, Ozaki T, Kitao M, Sugimura K. Diagnosis of pelvic endometriosis by magnetic resonance imaging using 'fat-saturation' technique. *Fertility and Sterility* 1994;**62**:973-7.

Takeuchi 2008 {published data only}

Takeuchi M, Matsuzaki K, Nishitani H. Susceptibility-weighted MRI of endometrioma: preliminary results. *American Journal of Roentgenology* 2008;**191**(5):1366-70.

Tammaa 2014 {published data only}

Tammaa A, Fritzer N, Strunk G, Krell A, Salzer H, Hudelist G. Learning curve for the detection of pouch of Douglas obliteration and deep infiltrating endometriosis of the rectum. *Human Reproduction* 2014;**29**(6):1199-204.

Tammaa 2015 {published data only}

Tammaa A, Fritzer N, Lozano P, Krell A, Salzer H, Salama M, et al. Interobserver agreement of non-invasive diagnosis of endometriosis by transvaginal sonography (TVS). *Ultrasound Obstetrics and Gynecology* 2015;**Epub Mar 12**:1-14.

Theodoridis 2009 {published data only}

Theodoridis TD, Zepiridis L, Mikos T, Grimbizis GF, Dinas K, Athanasiadis A, et al. Comparison of diagnostic accuracy of transvaginal ultrasound with laparoscopy in the management of patients with adnexal masses. *Archives of Gynecology and Obstetrics* 2009;**280**(5):767-73.

Valentini 2014 {published data only}

Valentini AL, Gui B, Micco M, Mingote MC, Ninivaggi V, Guido M, et al. How to improve MRI accuracy in detecting deep infiltrating colorectal endometriosis: MRI findings vs. laparoscopy and histopathology. *Radiologica Medica* 2014;**119**(5):291-7.

van Holsbeke 2010 {published data only}

Van Holsbeke C, Zhang J, Van Belle V, Paladini D, Guerriero S, Czekierdowski A, et al. Acoustic streaming cannot discriminate reliably between endometriomas and other types of adnexal lesion: a multicenter study of 633 adnexal masses. *Ultrasound in Obstetrics and Gynecology* 2010;**35**(3):349-53.

Vimercati 2012 {published data only}

Vimercati A, Achilarré MT, Scardapane A, Lorusso F, Ceci O, Mangiatordi G, et al. Accuracy of transvaginal sonography and contrast-enhanced magnetic resonance-colonography for the presurgical staging of deep infiltrating endometriosis. *Ultrasound in Obstetrics and Gynecology* 2012; Vol. 40, issue 5:592-603.

Volpi 1995 {published data only}

Volpi E, De Grandis T, Zuccaro G, La Vista A, Sismondi P. Role of transvaginal sonography in the detection of endometriomata. *Journal of Clinical Ultrasound* 1995;**23**:163-7.

Vrachnis 2012 {published data only}

Vrachnis N, Sifakis S, Samoli E, Kappou D, Pavlakis K, Iliodromiti Z, et al. Three-dimensional ultrasound and three-dimensional power Doppler improve the preoperative evaluation of complex benign ovarian lesions. *Clinical and Experimental Obstetrics and Gynecology* 2012;**39**(4):474-8.

Weerakiet 2000 {published data only}

Weerakiet S, Wongkularb A, Rochanawutanon M, Rojanasakul A. Transvaginal ultrasonography combined with pelvic examination in the diagnosis of ovarian endometrioma. *Journal of the Medical Association of Thailand* 2000;**83**(5):523-8.

Young 2013 {published data only}

Yong PJ, Sutton C, Suen M, Williams C. Endovaginal ultrasound-assisted pain mapping in endometriosis and chronic pelvic pain. *Journal of Obstetrics and Gynaecol* 2013;**33**(7):715-9.

Zanardi 2003 {published data only}

Zanardi R, Del Frate C, Zuiani C, Bazzocchi M. Staging of pelvic endometriosis based on MRI findings versus laparoscopic classification according to the American Fertility Society. *Abdominal Imaging* 2003;**28**:733-42.

Zawin 1989 {published data only}

Zawin M, McCarthy S, Scoutt L, Comite F. Endometriosis: appearance and detection at MR imaging. *Radiology* 1989;**171**(3):693-6.

Zykin 1981 {published data only}

Zykin BI. Echography in the diagnosis of genital endometriosis. *Sovetskaia Meditsina* 1981;**1**:51-4.

References to ongoing studies

NCT01939535 {unpublished data only}

Ongoing trial NCT01939535. Preoperative Staging of Endometriosis With MRI (IDEAL). Registered June 27, 2013.

NCT02233621 {unpublished data only}

Ongoing trial NCT02233621. Assessment of Performance of [18F]-FES for Endometriosis Diagnosis (ENDOTEP) [Evaluation Des Performances de la Tomographie Par Emission de Positons Avec la 16α-[18F]Fluoro-17β-estradiol ([18F]-FES) Pour le Diagnostic de l'Endometriose]. Registered June 2012.

NTR3738 {unpublished data only}

Ongoing trial NTR3738. Magnetic Resonance Imaging to Diagnose Endometriosis Using Ablavar® as Contrast Agent: A Feasibility Study. Registered 5-Dec-2012.

Additional references

ACOG Committee on Gynecology 2010

ACOG Committee on Adolescent Health Care. Practice Bulletin No. 114: Management of Endometriosis. *Obstetrics and Gynecology* 2010;**116**(1):223-36.

Adamson 2008

Adamson GD, Pasta DJ. Endometriosis Fertility Index (EFI): the new validated endometriosis staging system. *Fertility and Sterility* 2010;**94**(4):1609-15.

Almeida Filho 2008

Almeida Filho DP, Oliveira LJ, Amaral VF. Accuracy of laparoscopy for assessing patients with endometriosis. *Sao Paulo Medical Journal* 2008;**126**:305-8.

American Society for Reproductive Medicine 1997

American Society for Reproductive Medicine. Revised American Society for Reproductive Medicine classification of endometriosis: 1996. *Fertility and Sterility* 1997;**67**(5):817-21.

Ballard 2008

Ballard KD, Seaman HE, de Vries CS, Wright JT. Can symptomatology help in the diagnosis of endometriosis? Findings from a national case-control study - Part 1. *BJOG: an international journal of obstetrics and gynaecology* 2008;**115**(11):1382-91.

Batt 2003

Batt R, Mitwally MF. Endometriosis from thelarche to midteens: pathogenesis and prognosis, prevention and pedagogy. *Journal of Pediatric and Adolescent Gynecology* 2003;**16**:333-47.

Becker 2014

Becker CM, Laufer MR, Stratton P, Hummelshoj I, Missmer SA, Zondervan KT, et al. World Endometriosis Research Foundation Endometriosis Phenome and biobanking harmonization project: I. Surgical phenotype data collection in endometriosis research. *Fertility & Sterility* 2014;**102**(5):1213-22.

Bossuyt 2003

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *British Medical Journal* 2003;**326**(7379):41-4.

Bossuyt 2008

Bossuyt PM, Leeflang MM. Chapter 6: Developing Criteria for Including Studies. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* Version 0.4 [updated September 2008]. The Cochrane Collaboration, 2008.

Chapron 2003a

Chapron C, Fauconnier A, Vieira M, Barakat Dousset HB, Pansini V, Vacher-Lavenu MC, et al. Anatomical distribution of deeply infiltrating endometriosis: surgical implications and proposition for a classification. *Human Reproduction* 2003;**18**:157-61.

Chapron 2003b

Chapron C, Fauconnier A, Dubuisson JB, Barakat H, Vieira M, Bréart G. Deep infiltrating endometriosis: relation between severity of dysmenorrhea and extent of disease. *Human Reproduction* 2003;**18**:760-6.

Chapron 2003c

Chapron C, Cravello L, Chopin N, Kreiker G, Blanc B, Dubuisson JB. Complications during set-up procedures for laparoscopy in gynecology: open laparoscopy does not reduce the risk of major complications. *Acta Obstetrica et Gynecologica Scandinavica* 2003;**82**:1125-9.

de Vet 2008

de Vet HCW, Eisinga A, Riphagen II, Aertgeerts B, Pewsner D. Chapter 7: Searching for studies. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 [updated September 2008]. The Cochrane Collaboration, 2008.

Dmowski 1997

Dmowski WP, Lesniewicz R, Rana N, Pepping P, Noursalehi M. Changing trends in the diagnosis of endometriosis: a comparative study of women with pelvic endometriosis presenting with chronic pelvic pain or infertility. *Fertility and Sterility* 1997;**67**:238-43.

Duffy 2014

Duffy JMN, Arambage K, Correa FJS, Olive D, Farquhar C, Garry R, et al. Laparoscopic surgery for endometriosis. *Cochrane Database of Systematic Reviews* 2014;**4**:Art. No.: CD011031. [DOI: [10.1002/14651858.CD011031.pub2](https://doi.org/10.1002/14651858.CD011031.pub2)]

Dunselman 2014

Dunselman GA, Vermeulen N, Becker C, Calhaz-Jorge C, D'Hooghe T, De Bie B, et al. ESHRE guideline: management of women with endometriosis. *Human Reproduction* 2014;**29**(3):400-12.

Fauconnier 2005

Fauconnier A, Chapron C. Endometriosis and pelvic pain: epidemiological evidence of the relationship and implications. *Human Reproduction Update* 2005;**11**:595-606.

Frishman 2006

Frishman GN, Salak JR. Conservative surgical management of endometriosis in women with pelvic pain. *Journal of Minimally Invasive Gynecology* 2006;**13**:546-58.

Gao 2006

Gao X, Yeh YC, Outley J, Simon J, Botteman M, Spalding J. Health-related quality of life burden of women with endometriosis: a literature review. *Current Medical Research and Opinion* 2006;**22**:1787-97.

Garry 1997

Garry R. Laparoscopic excision of endometriosis: the treatment of choice?. *British Journal of Obstetrics and Gynaecology* 1997;**104**:513-5.

Giudice 2004

Giudice LC, Kao LC. Endometriosis. *Lancet* 2004;**364**:1789-99.

Greene 2009

Greene R, Stratton P, Cleary SD, Ballweg ML, Sinaii N. Diagnostic experience among 4,334 women reporting surgically diagnosed endometriosis. *Fertility and Sterility* 2009;**91**(1):32-9.

Guerriero 2015

Guerriero S, Ajossa S, Orozco R, Perniciano M, Jurado M, Melis GB, et al. Diagnostic accuracy of transvaginal ultrasound for diagnosis of deep endometriosis in the recto-sigmoid: a meta-analysis. *Ultrasound in Obstetrics & Gynecology* 2015;**Epub ahead of print**:1.

Guo 2009

Guo SW. Recurrence of endometriosis and its control. *Human Reproduction Update* 2009;**15**(4):441-61.

Guzick 1997

Guzick DS, Silliman NP, Adamson GD, Buttram VC Jr, Canis M, Malinak LR, et al. Prediction of pregnancy in infertile women based on the American Society for Reproductive Medicines revised classification of endometriosis. *Fertility and Sterility* 1997;**67**(5):822-9.

Halme 1984

Halme J, Hammond MG, Hulka JF, Raj SG, Talbert LM. Retrograde menstruation in healthy women and in patients with endometriosis. *Obstetrics and Gynecology* 1984;**64**(2):151-4.

Hudelist 2011b

Hudelist G, English J, Thomas AE, Tinelli A, Singer CF, Keckstein J. Diagnostic accuracy of transvaginal ultrasound for non-invasive diagnosis of bowel endometriosis: systematic review and meta-analysis. *Ultrasound in Obstetrics & Gynecology* 2011;**37**(3):257-63.

Johnson 2015

Johnson NP, et al. Consensus on the classification of endometriosis. *Human Reproduction* 2015;**in preparation/press**:1.

Johnson and Hummelshoj 2013

Johnson NP, Hummelshoj L. Consensus on current management of endometriosis. *Human Reproduction* 2013;**28**(6):1552-68.

Kennedy 2005

Kennedy S, Bergqvist A, Chapron C, D'Hooghe T, Dunselman G, Greb R, et al. ESHRE Special Interest Group for Endometriosis and Endometrium Guideline Development Group. ESHRE guideline for the diagnosis and treatment of endometriosis. *Human Reproduction* 2005;**20**(10):2698-704.

Kinkel 2006

Kinkel K, Frei KA, Balleyguier C, Chapron C. Diagnosis of endometriosis with imaging: a review. *European Radiology* 2006;**16**:285-98.

Koninckx 1991

Koninckx PR, Meuleman C, Demeyere S, Lesaffre E, Cornillie FJ. Suggestive evidence that pelvic endometriosis is a progressive

disease, whereas deeply infiltrating endometriosis is associated with pelvic pain. *Fertility and Sterility* 1991;**55**(4):759-65.

Koninckx and Martin 1994

Koninckx PR, Martin D. Treatment of deeply infiltrating endometriosis. *Current Opinion in Obstetrics and Gynecology* 1994;**6**(3):231-41.

Ling 1999

Ling F. Randomized controlled trial of depot leuprolide in patients with chronic pelvic pain and clinically suspected endometriosis. *Obstetrics and Gynecology* 1999;**93**:51-8.

Liu 2005

Liu A, Schisterman EF, Mazumdar M, Hu J. Power and sample size calculation of comparative diagnostic accuracy studies with multiple correlated test results. *Biometrical Journal* 2005;**47**(2):140-50.

Marchino 2005

Marchino GL, Gennarelli G, Enria R, Bongioanni F, Lipari G, Massobrio M. Diagnosis of pelvic endometriosis with use of macroscopic versus histologic findings. *Fertility and Sterility* 2005;**84**:12-5.

Martin 2001

Martin DC, Batt RE. Retrocervical, rectovaginal pouch, and rectovaginal septum endometriosis. *The Journal of the American Association of Gynecologic Laparoscopists* 2001;**8**(1):12-17.

Martin 2006

Martin DC. Applying STARD criteria to the laparoscopic identification of endometriosis (abstract). *Fertility and Sterility* 2006;**86**(Suppl 2):270.

Matsuzaki 2006

Matsuzaki S, Canis M, Pouly JL, Rabischong B, Botchorishvili R, Mage G. Relationship between delay of surgical diagnosis and severity of disease in patients with symptomatic deep infiltrating endometriosis. *Fertility and Sterility* 2006;**86**:1314-6.

McGraw-Hill Dictionary of Medicine 2006

Segen JC (author). McGraw-Hill Concise Dictionary of Modern Medicine. New York: The McGraw-Hill Companies, Inc, 2006.

Medeiros 2009

Medeiros LR, Rosa DD, Bozzetti MC, Fachel JM, Furness S, Garry R, et al. Laparoscopy versus laparotomy for benign ovarian tumour. *Cochrane Database of Systematic Reviews* 2009, Issue 2. [DOI: [10.1002/14651858.CD004751.pub3](https://doi.org/10.1002/14651858.CD004751.pub3)]

Medeiros 2014

Medeiros LR, Rosa MI, Silva BR, Reis ME, Simon CS, Dondossola ER, da Cunha Filho JS. Accuracy of magnetic resonance in deeply infiltrating endometriosis: a systematic review and meta-analysis. *Archives of gynecology and obstetrics* 2014;**Epub 2014/10/08**:1.

Moore 2002

Moore J, Copley S, Morris J, Lindsell D, Golding S, Kennedy S. A systematic review of the accuracy of ultrasound in the diagnosis

of endometriosis. *Ultrasound in Obstetrics and Gynecology* 2002;**20**:630-4.

Nyholt 2012

Nyholt DR, Low SK, Anderson CA, Painter JN, Uno S, Morris AP, et al. Genome-wide association meta-analysis identifies new endometriosis risk loci. *Nature Genetics* 2012;**44**(12):1355-9.

Redwine 2003

Redwine DB. Invisible' microscopic endometriosis: a review. *Gynecologic and Obstetric Investigation* 2003;**55**:63-7.

Rogers 2009

Rogers PA, D'Hooghe TM, Fazleabas A, Gargett CE, Giudice LC, Montgomery GW, et al. Priorities for Endometriosis Research: recommendations from an international consensus workshop. *Reproductive Sciences* 2009;**16**(4):335-46.

Rutjes 2005

Rutjes AWS, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PMM. Case-control and two-gate designs in diagnostic accuracy studies. *Clinical Chemistry* 2005;**51**(8):1335-41.

Saba 2014a

Saba L, Sulcis R, Melis GB, de Cecco CN, Laghi A, Piga M, et al. Endometriosis: the role of magnetic resonance imaging. *Acta Radiologica* 2014;**56**(3):355-67. [DOI: [10.1177/0284185114526086](https://doi.org/10.1177/0284185114526086)]

Sampson 1927

Sampson JA. Peritoneal endometriosis due to menstrual dissemination of endometrial tissue into the peritoneal cavity. *American Journal of Obstetrics and Gynecology* 1927;**14**:442-69.

Simoens 2012

Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. *Human Reproduction* 2012;**27**(5):1292-9.

Sinaii 2002

Sinaii N, Cleary SD, Ballweg ML, Nieman LK, Stratton P. High rates of autoimmune and endocrine disorders, fibromyalgia, chronic fatigue syndrome and atopic diseases among women with endometriosis: a survey analysis. *Human Reproduction* 2002;**17**(10):2715-24.

SOGC 2010

Society of Obstetricians Gynaecologists of Canada. Endometriosis: diagnosis and management. SOGC clinical practice guideline no. 244. *Journal of Obstetrics and Gynaecology Canada* 2010;**32**:S1-S28.

Somigliana 2006

Somigliana E, Vigano P, Parazzini F, Stoppelli S, Giambattista E, Vercellini P. Association between endometriosis and cancer: a comprehensive review and a critical analysis of clinical and epidemiological evidence. *Gynecologic Oncology* 2006;**101**(2):331-41.

Spaczynski 2003

Spaczynski RZ, Duleba AJ. Diagnosis of endometriosis. *Seminars in Reproductive Medicine* 2003;**21**:193-208.

Stegmann 2008

Stegmann BJ, Sinaii N, Liu S, Segars J, Merino M, Nieman LK, Stratton P. Using location, color, size, and depth to characterize and identify endometriosis lesions in a cohort of 133 women. *Fertility and Sterility* 2008;**89**:1632-6.

The Gale Encyclopedia of Medicine 2008

Olendorf D (Editor), Jeryan C (Editor), Boyden K (Editor), Gale Group (Corporate Author). The Gale Encyclopedia of Medicine (5 volume set). Farmington Hills, MI: The Gale Group, Inc, 2008.

Vercellini 1996

Vercellini P, Trespidi L, De Giorgi O, Cortesi I, Parazzini F, Crosignani GP. Endometriosis and pelvic pain: relation to disease stage and localization. *Fertility and Sterility* 1996;**65**:299-304.

Vigano 2004

Vigano P, Parazzini F, Somigliana E, Vercellini P. Endometriosis: epidemiology and aetiological factors. *Best Practice & Research: Clinical Obstetrics & Gynaecology* 2004;**18**:177-200.

Vitonis 2014

Vitonis AF, Vincent K, Rahmioglu N, Fassbender A, Buck Louis G, Hummelshoj L, et al. World Endometriosis Research Foundation Endometriosis Phenome and biobanking harmonization project: II. Clinical and covariate phenotype data collection in endometriosis research. *Fertility & Sterility* 2014;**102**(5):1223-32.

Whiting 2005

Whiting PF, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology* 2005;**5**:19.

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. the QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36.

Wykes 2004

Wykes CB, Clark TJ, Khan KS. Accuracy of laparoscopy in the diagnosis of endometriosis: a systematic quantitative review. *BJOG - an International Journal of Obstetrics and Gynaecology* 2004;**111**:1204-12.

Yeung 2009

Yeung PP Jr, Shwayder J, Pasic RP. Laparoscopic management of endometriosis: comprehensive review of best evidence. *Journal of Minimally Invasive Gynecology* 2009;**16**:269-81.

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Abrao 2007

Study characteristics	
Patient sampling	<p>Primary objectives: to evaluate the capacity of clinical examination (digital vaginal examination), trans-vaginal ultrasonography (TVUS) and pelvic magnetic resonance imaging (MRI) in patients with clinical suspicion of endometriosis in the rectosigmoid and/or retrocervical region and to compare accuracy of these techniques</p> <p>Study population: patients with clinically suspected endometriosis</p> <p>Selection criteria: exclusion criteria: virgin or individual with any type of genital malformation that made physical examination or transvaginal ultrasonography impossible; unable to tolerate MRI</p> <p>Study design: prospective cross-sectional; consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 53/104, deep dyspareunia 66/104, acyclical pelvic pain 17/104, infertility 55/104, cyclical bowel symptoms (pain/bleeding) 59/104, cyclical urinary symptoms 14/104</p> <p>Age: mean 33.8 ± 6.1 years, range 18 to 45 years</p> <p>Number enrolled: 104 women</p> <p>Number available for analysis: 104 women</p> <p>Setting: tertiary university hospital, referral centre for endometriosis, São Paulo University</p> <p>Place of study: São Paulo, Brazil</p> <p>Period of study: August 2004 to October 2006</p>

Abrao 2007 (Continued)

Language: English

Index tests	<p>Index test: <u>TVUS</u>; <u>MRI</u> (T1/T2-w)</p> <p>Description of positive case definition by index test as reported: <u>TVUS</u> - deep retrocervical endometriosis defined as thick blocks of tissue, nodular formations or irregular shaped, hypoechoic, retractable masses in USL, POD and/or vagina; bowel involvement established as a long, nodular, predominantly solid, hypoechogenic lesion adhered to the wall of the intestinal loop; each examination interpreted in real time; <u>MRI</u> - retrocervical endometriosis defined as USL of irregular thickness or as retractable nodules with spiky edges with low signal in T1/T2-w MR images, with or without cysts, unilateral or bilateral, or as nodules or irregular thick blocks of tissue with low signal in T1/T2-w images, situated posterior to the cervix, near the vaginal dome; bowel involvement identified as retractable nodular formations adhered to the bowel wall, with a strong hyposignal in T2 demonstrating delayed gadolinium enhancement, identified on anterior wall of rectum, rectosigmoid junction, sigmoid colon, caecum and intestinal loops; image quality good in all cases</p> <p>Examiners: TVUS and MRI carried out independently by a single examiner who was blinded to participants' clinical data and to results of other imaging; level of expertise not reported</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: posterior DIE (rectosigmoid and retrocervical area) - separate anatomical sites</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 98/104 (91%), DIE 63/104 (61%)</p> <p>Reference standard: laparoscopy 104/104 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: visual inspection + histological confirmation: criteria not specified; surgical procedure not described</p> <p>Examiners: number or level of expertise of surgeons or pathologists not reported; not blinded to index test result - "the decision regarding surgical procedure to be carried out was based on both clinical exam and imaging results"</p>
Flow and timing	<p>Time interval between index test and reference standard: within 3 months</p> <p>Withdrawals: none reported</p>
Comparative	
Key conclusions by the authors	<p>TVUS had better sensitivity, specificity, PPV, NPV and accuracy in cases of deep retrocervical and rectosigmoid endometriosis when compared with MRI and digital vaginal examination, confirming that it is an important preoperative examination for the definition of surgical strategies</p>
Conflict of interests	Not reported
Notes	<p>Reported accuracy estimates for TVUS and MRI for diagnosis of specific sites of DIE confirmed as accurate</p> <p>No data available on the accuracy of index tests for overall DIE</p> <p>Accuracy estimates of pelvic examination and comparisons of pelvic examination with index tests presented - not included in this review</p> <p>Possible overlap of MRI data with another study from the same group, Chamie 2009 (study period November 2005 to July 2007); unable to clarify with study authors; therefore, results of both studies are included in this review</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

Abrao 2007 (Continued)

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled? Yes

Did the study avoid inappropriate exclusions? Yes

Was a 'two-gate' design avoided? Yes

Low
Low
DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test? Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice? Yes

Low
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? No

High
Low
DOMAIN 4: Flow and Timing

Abrao 2007 (Continued)

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Ascher 1995
Study characteristics

Patient sampling **Primary objectives:** to assess conventional spin echo (CSE) alone and in combination with T1-w fat-suppressed (TIFS) and gadolinium-enhanced TIFS (Gd-TIFS) spin-echo techniques for detection of endometriosis with laparoscopy or laparotomy as a “gold standard”

Study population: women with clinically suspected endometriosis who were scheduled for surgery

Selection criteria: not specified

Study design: prospective observational; non-consecutive enrolment

Patient characteristics and setting

Clinical presentation: not specified

Age: mean 34.1 years, range 21 to 46 years

Number enrolled: 38 women

Number available for analysis: 31 women

Setting: not specified

Place of study: USA

Period of study: 11-month period, dates not specified

Language: English

Index tests

Index test: MRI 3 types (T1/T2-w (CSE); T1/T2-w + fat-suppressed (CSE/TIFS); T1/T2-w + fat-suppressed + Gd (CSE/TIFS/Gd-TIFS))

Description of positive case definition by index test as reported: endometriomas diagnosed by published criteria (referenced to Nishimura et al., 1987; Togashi et al., 1991; and Sugimura et al., 1993) and described; diagnosis not applied to huge lesions or lesions with septations or solid components; implants diagnosed as ill-defined peritoneum-based regions of enhancement without discrete masses on contrast-enhanced images and not attempted with non-contrast images; image quality good in 26 cases. fair in 4 and poor in 1

Examiners: MR images prospectively evaluated by 2 radiologists experienced in pelvic MRI; readers aware of clinical suspicion of endometriosis

Ascher 1995 (Continued)

	<p>Interobserver variability: the 2 separate image readers agreed in 27 of 31 cases; consensus reading reached in remaining 4 cases</p>
Target condition and reference standard(s)	<p>Target condition: pelvic endometriosis</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 21/31 (67.7%)</p> <p>Reference standard: laparoscopy 24/31 (77.4%), laparotomy 7/31 (22.6%)</p> <p>Description of positive case definition by reference test as reported: at surgery, pelvis examined in 5 regions: right and left adnexae, surface of uterus, cul-de-sac and peritoneum; diagnosis made by direct visualisation of endometriotic implants</p> <p>Examiners: numbers or level of expertise of surgeons not reported; unclear whether blinded to results of index tests</p>
Flow and timing	<p>Time interval between index test and reference standard: within 12 weeks</p> <p>Withdrawals: 7 enrolled participants (18%) excluded for the following: incomplete studies and/or did not undergo surgery</p>
Comparative	
Key conclusions by the authors	<p>In summary, sensitivity, specificity and accuracy of MR imaging for investigation of patients with suspected endometriosis make it a moderately useful modality. Significantly improved ability to detect small endometriomas when TIFS or TIFS/Gd-T1FS is combined with CSE images. For implant detection, the role of contrast enhancement less clear; may provide information about some endometriosis implants, but low sensitivity.</p>
Conflict of interests	Not reported
Notes	<p>Reported accuracy estimates for different MRI modalities for diagnosis of pelvic endometriosis confirmed as accurate.</p> <p>Unclear whether peritoneal endometriosis was assessed with conventional or fat-suppressed MRI without contrast, given the statement "implant detection was not attempted with non-contrast imaging".</p> <p>Separate accuracy estimates also presented for peritoneal implants as well as for large and small endometriomas - not included in this review (data for numbers of lesions, not numbers of participants).</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			

Ascher 1995 (Continued)

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes	
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes	
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes	
	Low	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Unclear	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
	High	

Bazot 2009
Study characteristics

Patient sampling **Primary objectives:** to compare the value of physical examination, TVS, RES and MRI for the diagnosis of different locations of DIE

Bazot 2009 (Continued)

Study population: women referred with clinical evidence of pelvic endometriosis

Selection criteria: not specified

Study design: longitudinal; consecutive enrolment

Patient characteristics and setting

Clinical presentation: dysmenorrhoea 79/92, dyspareunia 63/92, dyschezia 32/92, dysuria 3/92, infertility 21/92; history of surgery for endometriosis 31/92

Age: median age 31.8 years, range 20 to 50 years

Number enrolled: 92 women

Number available for analysis: 92 women

Setting: tertiary care Tenon Hospital, referral centre for endometriosis and Surgical Centre Trocadero

Place of study: Paris, France

Period of study: April 2000 to May 2005

Language: English

Index tests

Index test: TVUS (TVS); TRUS (RES); MRI (T1/T2-w + fat-suppressed/Gd)

Description of positive case definition by index test as reported: all examinations performed and interpreted in real time and videotaped for review; all potential locations of endometriosis examined; diagnostic criteria provided for each test for each anatomical site of endometriosis with reference to published criteria ([Bazot 2003](#); [Bazot 2004a](#); [Bazot 2004b](#)).

Examiners: all techniques interpreted independently and blindly by different physicians. TVS: all scans performed by a single radiologist with extensive experience in gynaecological imaging. RES: each examination interpreted in real time by the same gastroenterologist with 5 years' experience in endometriosis. MRI: each examination interpreted according to a standardised protocol, retrospectively by 1 radiologist with 2 years' experience in gynaecological imaging. Readers informed of women's clinical history and symptoms but blinded to results of physical and previous imaging examinations

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: DIE: separate anatomical sites; ovarian endometriosis

Prevalence of target condition in the sample: DIE 90/92 (97.8%); ovarian endometriosis 36/92 (39.1%)

Reference standard: laparoscopy 79/92 (85.9%), laparotomy 13/92 (14.1%) + histopathology

Description of positive case definition by reference test as reported: all locations of endometriosis recorded on surgical reports. Histological criteria described and referenced to primary source (Clement 2002); DIE diagnosed if clearly visualised lesions, but fibrosis/smooth muscle cell on histology or another histologically proven site of endometriosis was found when lesion was not biopsied, or if complete cul-de-sac obliteration secondary to endometriosis was observed; surgical procedure not described

Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinded to results of index tests

Flow and timing

Time interval between index test and reference standard: < 12 months (personal communication with study authors)

Withdrawals: none reported

Comparative

Bazot 2009 (Continued)

Key conclusions by the authors MRI provides a more reliable map of DIE than physical examination, TVS or RES. In women with chronic pelvic pain suggestive of pelvic endometriosis, TVS should remain the first-line technique examination, although normal TVS findings do not rule out the diagnosis. Hence, MRI should be used to examine symptomatic women before surgery. Use of RES should be restricted to cases in which a discrepancy is found between physical examination and first-line imaging techniques

Conflict of interests Study authors declared no conflict of interest

Notes Reported accuracy estimates for TVS for diagnosis at different sites of posterior DIE confirmed as accurate; accuracy estimates for endometrioma not presented by study authors and calculated on the basis of narrative data from text

No data available for calculating accuracy estimates for overall DIE for TVUS and RES and for anterior DIE for all tests

Accuracy estimates for pelvic examination also presented - not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be avail-	Yes		

Bazot 2009 (Continued)

able when the test is used
in practice?

	Low	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Bazot 2013
Study characteristics

Patient sampling	<p>Primary objectives: to compare overall image quality and diagnostic accuracy of multi-planar 2-dimensional (2D) fast spin-echo (FSE) T2-w and 3-dimensional (3D) coronal single-slab FSE T2-w magnetic resonance imaging (MRI) sequence for evaluation of deep infiltrating endometriosis (DIE)</p> <p>Study population: patients referred for pelvic MRI because of clinical suspicion of endometriosis</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational, consecutive enrolment of patients presented to imaging department</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea, deep dyspareunia, dyschezia, dysuria or infertility</p> <p>Age: median age 34 years, range 24 to 46 years</p> <p>Number enrolled: 110 women</p> <p>Number available for analysis: 23 women</p> <p>Setting: tertiary care hospital, Tenon Hospital, referral centre for endometriosis</p>

Bazot 2013 (Continued)

Place of study: Paris, France

Period of study: February 2010 to May 2010

Language: English

Index tests	<p>Index test: MRI 2 types: 2-dimensional fast spin echo T2-w (2D FSE T2-w MRI); 3-dimensional fast spin echo T2-w MRI (3D FSE T2-w MRI)</p> <p>Description of positive case definition by index test as reported: diagnostic criteria mentioned and referenced to a primary source (Kinkel et al., 1999; Bazot et al., 2004; Kataoka et al., 2005). Readers asked to determine overall image quality and presence or absence of DIE</p> <p>Examiners: images independently analysed by 2 radiologists with different degrees of experience in female MRI (1 reader with > 20 years' experience; second reader a junior radiologist). Both readers blinded to clinical and ultrasonographic findings</p> <p>Interobserver variability: poor interobserver agreement for assessment of DIE found for USL endometriosis. For all locations of endometriosis, high intraobserver agreement observed for an experienced reader; low intraobserver agreement for USL, rectosigmoid and POD obliteration for junior reader</p>
Target condition and reference standard(s)	<p>Target condition: posterior DIE: overall and separate anatomical sites</p> <p>Prevalence of target condition in the sample: pelvic endometriosis in 20/23 (87%); DIE 18/23 (78%); specific locations of DIE: USL 17/23 (74%), rectosigmoid 13/23 (57%), vaginal 5/23 (22%), bladder 1/23 (4%)</p> <p>Reference standard: laparoscopy (n = 20), laparotomy (n = 3) + histopathology.</p> <p>Description of positive case definition by reference test as reported: all locations of endometriosis recorded in surgical reports; reference to sources for histological criteria; surgical procedure not described</p> <p>Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinded to results of index test</p>
Flow and timing	<p>Time interval between index test and reference standard: < 12 months (communication with study authors)</p> <p>Withdrawals: 87/110 (79%) women did not undergo surgery and were excluded from final analysis, reason not explained</p>
Comparative	
Key conclusions by the authors	<p>Accuracy of 3D MRI yields accuracy not significantly different from accuracy of 2D FSE T2-w MRI in diagnosis of DIE locations. However, despite significant time savings, 3D MRI cannot replace routine 2D MRI sequences because poorer imaging quality results from significant intraobserver and interobserver variability</p>
Conflict of interests	Not reported
Notes	<p>Reported accuracy estimates for index tests for diagnosis of overall DIE and different sites of posterior DIE confirmed as accurate</p> <p>Accuracy estimates presented separately by study authors for each reader. Only data from R1 (experienced reader) reader presented in this review</p> <p>Overall image quality for 2 MRI techniques and detailed assessment of interobserver and intraobserver variability presented by study authors - not presented in this review</p> <p>Only 21% of participants underwent surgery, hence high selection bias</p>

Bazot 2013 (Continued)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low

Bazot 2013 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
High	

Bergamini 2010
Study characteristics

Patient sampling	<p>Primary objectives: to evaluate accuracy of transrectal sonography (TRS) and a new technique, transvaginal sonography with water contrast in the rectum (RWC-TVS), in the diagnosis of rectosigmoid endometriosis, and accuracy of barium enema (BE) and RWC-TVS in detection of intestinal stenosis due to endometriosis; to describe our experience with a new diagnostic imaging approach for preoperative assessment of intestinal endometriosis including identification of cases that are candidates for segmental bowel resection</p> <p>Study population: women scheduled for surgery because of signs and symptoms of severe posterior deep infiltrating endometriosis</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, multi-centre, observational; consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dyspareunia and/or catamenial rectal pain 61/61, history of intermittent bowel obstruction 4/61, nulliparous 11/61, history of surgery for endometriosis 19/61</p> <p>Age: mean age 33.1 years, range 28 to 37 years</p> <p>Number enrolled: 61 women</p> <p>Number available for analysis: 61 women</p> <p>Setting: University Hospitals of Verona and Varese, referral centres for endometriosis treatment</p> <p>Place of study: Verona and Varese, Italy</p> <p>Period of study: January 2008 to February 2009</p> <p>Language: English</p>
Index tests	<p>Index tests: <u>TRUS</u> (TRS); <u>TVUS</u> (RWC-TVS)</p> <p>Description of positive case definition by index test as reported: uterine cervix, parametria, uterosacral ligaments and vaginal and rectal walls up to the rectosigmoid junction evaluated; images of endometriotic lesions obtained with both techniques and recorded; definition of endometriotic lesions not prespecified (example images provided)</p> <p>Examiners: all scans performed by the same operator (gynaecologist), who had extensive experience in ultrasonographic diagnosis of endometriosis. Operator blinded with respect to other diag-</p>

Bergamini 2010 (Continued)

nostic findings; unclear whether operator was aware of the results of an additional index test (same operator, different test times)

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: rectosigmoid endometriosis

Prevalence of target condition in the sample: pelvic endometriosis 58/61 (95%), rectosigmoid endometriosis 51/61 (84%)

Reference standard: laparoscopy 57/61 (93.4%), laparotomy 4/61 (6.6%) + histopathology

Description of positive case definition by reference test as reported: direct visualisation ± histological examination; criteria not specified; surgical procedure described

Examiners: numbers or level of expertise of surgeons or pathologists not reported; no blinding to results of index tests; "segmental bowel resection was based on radiographic criteria"

Flow and timing

Time interval between index test and reference standard: not specified, but statement 'subsequently, all women underwent surgical treatment' allows one to assume that the interval was reasonably short

Withdrawals: none reported

Comparative

Key conclusions by the authors

RWC-TVS is a new, simple technique for single-step and accurate preoperative assessment of rectosigmoid endometriosis

Conflict of interests

Not reported

Notes

Reported accuracy estimates for TRS and RWC-TVS for diagnosis of rectosigmoid endometriosis confirmed as accurate

Accuracy estimates for BE and RWC-TVS for diagnosis of intestinal lumen stenosis also presented - not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Bergamini 2010 *(Continued)*

Did the study provide a clear pre-specified definition of what was considered to be a “positive” result of index test?	No
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear
High Low	
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
High Low	
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Unclear	

Biscaldi 2007
Study characteristics

Patient sampling	<p>Primary objectives: to investigate the efficacy of multi-slice computed tomography combined with colon distension by water enteroclysis (MSCTe) in the diagnosis of bowel endometriosis</p> <p>Study population: women who had both typical symptoms caused by pelvic endometriosis and gastrointestinal symptoms suggestive of colorectal endometriosis</p> <p>Selection criteria: not specified</p>
------------------	---

Biscaldi 2007 (Continued)

	Study design: prospective, observational; non-consecutive enrolment
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 87/98, dyspareunia 73/98, chronic pelvic pain 48/98, infertility 23/98, diarrhoea 20/98, constipation 12/98, bloating 5/98; previous surgery for endometriosis 37/98, previous medical treatment: oral contraceptive pill 81/98, GnRH analogues 40/98, norethisterone acetate 7/98, letrozole 2/98; no patients with previous bowel surgery other than appendicectomy</p> <p>Age: median age 34 years, range 20 to 53 years</p> <p>Number enrolled: 98 women</p> <p>Number available for analysis: 98 women</p> <p>Setting: tertiary care university hospital, San Martino Hospital, referral centre for endometriosis, Galliera Hospital</p> <p>Place of study: Genoa, Italy</p> <p>Period of study: January 2004 to December 2005</p> <p>Language: English</p>
Index tests	<p>Index test: MDCT-e (MSCTe)</p> <p>Description of positive case definition by index test as reported: MSCT criterion for diagnosis of bowel endometriosis (sigmoid, rectum, caecum, ileum) was presence of solid nodules with positive enhancement, contiguous or penetrating the thickened colonic wall; characteristics of involvement of different layers of bowel wall described</p> <p>Examiners: images independently reviewed by 2 observers; level of expertise not reported; radiologists not aware of clinical findings and patient history, knowing only that bowel endometriosis was suspected</p> <p>Interobserver variability: not presented; disagreement between observers resolved by consensus in a joined session</p>
Target condition and reference standard(s)	<p>Target condition: bowel endometriosis</p> <p>Prevalence of target condition in the sample: bowel endometriosis 76/98 (77.5%)</p> <p>Reference standard: laparoscopy 98/98 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: anatomical distribution of bowel endometriotic lesions recorded during surgery and histologically evaluated, reference to a source of histological criteria; surgical procedure described in details</p> <p>Examiners: all surgical procedures performed by a team of gynaecological and colorectal surgeons with extensive experience in the treatment of bowel endometriosis; unclear whether blinded to results of index test</p>
Flow and timing	<p>Time interval between index test and reference standard: within 20 days</p> <p>Withdrawals: none reported</p>
Comparative	
Key conclusions by the authors	MSCT combined with colon retrograde distension effective in the diagnosis of bowel endometriosis
Conflict of interests	Not reported

Biscaldi 2007 (Continued)

Notes

Reported accuracy estimates for MSCTe for diagnosis of bowel endometriosis confirmed as accurate

Accuracy estimates for MSCTe to estimate degree of bowel involvement, diameter of lesion and correlation with histopathology also presented - not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low

Biscaldi 2007 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Biscaldi 2014
Study characteristics

Patient sampling	<p>Primary objectives: to compare the accuracy of multi-detector computerised tomography enema (MDCT-e) and magnetic resonance enema (MRI-e) in determining the presence of rectal and sigmoid endometriotic nodules</p> <p>Study population: patients referred to (our) endometriosis centre</p> <p>Selection criteria: Inclusion criteria: reproductive age, suspicion of deep pelvic endometriosis on the basis of symptoms and vaginal examination, gastrointestinal symptoms that might be caused by rectosigmoid endometriosis. Exclusion criteria: previous bilateral ovariectomy, previous radiological exams of the bowel requiring contrast media, previous bowel surgery (except appendectomy), history of intolerance to iodinated contrast media, renal or hepatic failure, contraindications to MR examination, psychiatric disorders</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 185/260, dyspareunia 157/260, chronic pelvic pain 142/260, infertility 54/260, diarrhoea 57/260, constipation 85/260, bloating 122/260, dyschezia 130/260; previous surgery for endometriosis 113/260, previous medical treatment: oral contraceptive pill 79/260, contraceptive vaginal ring 14/260</p> <p>Age: mean 32.6 ± 4.3 years</p> <p>Number enrolled: 260 women</p> <p>Number available for analysis: 260 women</p> <p>Setting: tertiary care university hospital, San Martino Hospital, referral centre for endometriosis, Galliera Hospital</p> <p>Place of study: Genoa, Italy</p> <p>Period of study: not specified</p> <p>Language: English</p>
Index tests	<p>Index test: MDCT-e; MRI jelly method (MRI-e)</p> <p>Description of positive case definition by index test as reported: MDCT criterion for diagnosis of bowel endometriosis was presence of solid nodules contiguous or penetrating the thickened colonic wall. Infiltration of muscularis propria diagnosed when fat plane between nodule and bowel disappears</p>

Biscaldi 2014 (Continued)

with positive enhancement, and nodule penetrates intestinal wall from outside, licks inner surface and bulges toward mucosa. **MRI-e** criterion: visible penetration of endometriotic nodules in the intestinal wall; nodules defined as solid masses outside the sigmoid or rectal wall, frequently with hypointense signal due to their fibrous nature

Examiners: 2 radiologists blindly reviewed images at a PACS workstation; they were not aware of clinical findings and patient history, knowing only that the presence of bowel endometriosis was clinically suspected; level of expertise not reported

Interobserver variability: not presented; disagreement between observers resolved by consensus in a joined session

Target condition and reference standard(s)

Target condition: RS endometriosis

Prevalence of target condition in the sample: bowel endometriosis 176/260 (67.7%)

Reference standard: laparoscopy 260/260 (100%) + histopathology

Description of positive case definition by reference test as reported: bowel endometriosis defined as endometriotic lesions infiltrating at least the muscularis propria of the intestinal wall; sigmoid colon and rectum systematically examined to verify the presence of endometriotic lesions; all surgical specimens histologically evaluated, reference to the source (Remorgida et al., 2005); surgical procedure described

Examiners: all surgical procedures performed by a team of gynaecological and colorectal surgeons with extensive experience in the treatment of bowel endometriosis; surgeons aware of results of index tests

Flow and timing

Time interval between index test and reference standard: within 1 month

Withdrawals: none reported

Comparative

Key conclusions by the authors

In conclusion, both MDCT-e and MRI-e are accurate in the diagnosis of rectosigmoid endometriosis. MDCT-e has the disadvantage of using ionising radiation and iodinated contrast medium in a population of women of reproductive age. MRI-e more tolerable than MDCT-e

Conflict of interests

Not reported

Notes

Reported accuracy estimates for MDCT-e and MRI-e for diagnosis of RS endometriosis confirmed as accurate

Agreement between index test and histopathology for different sizes of lesions presented - not included in this review

Methodological quality

Item

Authors' judgement

Risk of bias

Applicability concerns

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?

No

Did the study avoid inappropriate exclusions?

Yes

Biscaldi 2014 (Continued)

Was a 'two-gate' design avoided? Yes

	High	Low
DOMAIN 2: Index Test Any test		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes	
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes	
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes	
	Low	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	No	
	High	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	

Biscaldi 2014 (Continued)

Low

Chamie 2009a

Study characteristics

Patient sampling	<p>Primary objectives: to evaluate the accuracy of magnetic resonance imaging (MRI) findings for diagnosis of deeply infiltrating endometriosis (DIE) at multiple sites, such as retrocervical space, rectosigmoid, bladder, ureters and vagina</p> <p>Study population: women who had a history and findings of a physical exam consistent with endometriosis</p> <p>Selection criteria: Inclusion criteria: symptoms consistent with endometriosis, such as pelvic pain, dysmenorrhoea, deep dyspareunia, acyclical pelvic pain, dyschezia and infertility; pelvic examination revealing thickening of posterior cul-de-sac and/or nodules; transvaginal ultrasound results showing ovarian cysts with thickened low-amplitude echoes; no previous pelvic surgery for endometriosis</p> <p>Study design: prospective, cross-sectional; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 89/92, dyspareunia 54/92, acyclical pain 72/92, dysuria 8/92, dyschezia 44/92, infertility 40/92; painful palpable nodules on examination 58/92</p> <p>Age: mean 33 years, range 20 to 52 years</p> <p>Number enrolled: 92 women</p> <p>Number available for analysis: 92 women</p> <p>Setting: tertiary university hospital, referral centre for endometriosis, São Paulo University</p> <p>Place of study: São Paulo, Brazil</p> <p>Period of study: November 2005 to July 2007</p> <p>Language: English</p>
Index tests	<p>Index test: MRI (T1/T2-w + fat-suppressed/Gd)</p> <p>Description of positive case definition by index test as reported: DIE diagnosed according to signal intensity and morphological abnormalities as previously described (referenced to Bazot et al., 2004); assessed sites included retrocervical region, rectosigmoid, bladder, ureters and vagina</p> <p>Examiners: MR images analysed prospectively by 2 radiologists (same examiners) who were blinded to each patient's history, physical findings and ultrasound results; level of expertise not reported</p> <p>Interobserver variability: not provided; MRI findings recorded as a consensus between the 2 radiologists</p>
Target condition and reference standard(s)	<p>Target condition: DIE - separate anatomical sites</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 92/92 (100%), DIE 77/92 (83.7%)</p> <p>Reference standard: laparoscopy 92/92 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: criteria not specified; surgical procedure reported as "extensive laparoscopic surgery" but not described</p>

Chamie 2009a (Continued)

Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinded to results of index test

Flow and timing	<p>Time interval between index test and reference standard: not specified, but statement "all the patients that were included underwent pelvic MRI before extensive laparoscopic surgery" allows one to assume that the interval was reasonably short</p> <p>Withdrawals: none reported</p>
Comparative	
Key conclusions by the authors	Present findings indicate perioperative MRI as excellent tool to provide reasonably accurate mapping of multiple sites of pelvic endometriosis
Conflict of interests	Not reported
Notes	<p>Reported accuracy estimates for MRI for diagnosis of different sites of DIE confirmed as accurate</p> <p>Accuracy estimates for bladder and ureteric endometriosis reported by study authors but not presented in the review because these target conditions were not assessed</p> <p>Possible overlap of MRI data with another study from the same group (Abrao 2007a (study period August 2004 to October 2006)): unable to clarify with study authors; therefore, results of both studies included in this review</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		

Chamie 2009a (Continued)

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes	
		<div>Low</div> <div>Low</div>
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
		<div>Unclear</div> <div>Low</div>
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Unclear	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
		<div>Unclear</div>

Dessole 2003

Study characteristics		
Patient sampling	<p>Primary objectives: to assess the accuracy of transvaginal ultrasonography and of sonovaginography for detection and location and extension assessment of rectovaginal endometriotic lesions, as well as to compare patient compliance between procedures</p> <p>Study population: women scheduled for laparotomy or laparoscopy because rectovaginal endometriosis is suspected on the basis of patient history and clinical examination</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>	
Patient characteristics and setting	<p>Clinical presentation: chronic pelvic pain, dysmenorrhoea or dyspareunia 38/46, infertility 20/46, gastrointestinal disorders 7/46, urinary disorders 6/46; endometriotic lesion detected on gynaecological examination 8/46; no patients had undergone surgical pelvic procedure before entering the study</p> <p>Age: mean 30.3 ± 4.2 years</p>	

Dessole 2003 (Continued)

Number enrolled: 46 women

Number available for analysis: 46 women

Setting: University Hospital, University of Sassari

Place of study: Sassari, Italy

Period of study: January 2000 to October 2001

Language: English

Index tests	<p>Index test: <u>TVUS</u> (transvaginal ultrasonography); <u>sonovaginography</u></p> <p>Description of positive case definition by index test as reported: <u>TVUS</u> - operator obtained longitudinal and transversal scans of the uterus, with particular attention given to rectovaginal septum for detection of endometriotic lesions - criteria not specified; <u>sonovaginography</u> - endometriotic lesions detected as hypoechoic, irregular structures at the level of the vaginal wall; often infiltrated surrounding structures (vesicovaginal or rectovaginal septum, rectal wall, Douglas pouch and USL); unclear whether prespecified criteria or description of findings</p> <p>Examiners: numbers of examiners, level of expertise and blinding to clinical data not reported</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: posterior DIE (rectovaginal endometriosis)</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 40/46 (87%), rectovaginal endometriosis 32/46 (69.5%), peritoneal endometriosis 8/46 (17.4%)</p> <p>Reference standard: laparoscopy 20/46 (43.5%), laparotomy 26/46 (56.5%) + histopathology</p> <p>Description of positive case definition by reference test as reported: direct visualisation ± histological examination; criteria not specified; surgical procedure described</p> <p>Examiners: numbers or level of expertise of surgeons or pathologists not reported; no blinding to results of index test: "The kind of operation was chosen on the basis of clinical and ultrasonographic findings"</p>
Flow and timing	<p>Time interval between index test and reference standard: not specified, but statement "before surgery, all the women underwent transvaginal ultrasonography and then sonovaginography" allows one to assume that interval was reasonably short</p> <p>Withdrawals: none reported</p>
Comparative	
Key conclusions by the authors	Sonovaginography is a reliable and simple method for assessment of rectovaginal endometriosis; it provides information on location, extension and infiltration of lesions - important factors in selecting the kind of surgery
Conflict of interests	Not reported
Notes	Presented accuracy estimates for transvaginal ultrasound and sonovaginography for diagnosis of RVS endometriosis confirmed as accurate

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Dessole 2003 (Continued)

Was a consecutive or random sample of patients enrolled?	No
Did the study avoid inappropriate exclusions?	No
Was a 'two-gate' design avoided?	Yes
<div>HighLow</div>	

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	No
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear
<div>HighLow</div>	

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
<div>HighLow</div>	

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes

Dessole 2003 (Continued)

Were all patients included in the analysis? Yes

Unclear

Eskenazi 2001
Study characteristics

Patient sampling	<p>Primary objectives: to determine whether surgical diagnosis of endometriosis can be predicted using common non-invasive tools including medical history, symptom report, pelvic examination and transvaginal ultrasound. We develop a predictive decision tree based on 1 sample of women who are about to undergo laparoscopy (study sample) and test the utility of this decision tree on a different sample of women who underwent laparoscopy (test sample)</p> <p>Study population: women scheduled to undergo laparoscopy or laparotomy for pelvic pain, infertility, tubal ligation or adnexal/uterine masses</p> <p>Selection criteria: exclusion criteria: acute conditions such as ectopic pregnancy, evaluation of endometrial or ovarian cancer, treatment of already diagnosed endometriosis</p> <p>Study design: prospective, observational; non-consecutive enrolment (study sample); retrospective record review (test sample)</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 40/90, pelvic pain 20/90, dyspareunia 20/90, infertility 12/90, abnormal pelvic examination 42/90; indications for surgery including pelvic pain 21%, infertility 13%, ovarian cysts 30%, fibroids 28%, suspected endometriosis 16%, tubal ligation 6.7%; nulliparous 42/90, nulligravid 33/90, current oral contraceptive users 4/90</p> <p>Age: mean 35.7 ± 7.2 years, range 20 to 49 years</p> <p>Number enrolled: 90 women (study sample); 120 women (test sample)</p> <p>Number available for analysis: 90 women – only 'study sample' arm included in current analysis; 'test sample' excluded for retrospective design</p> <p>Setting: Hospital of Desio (study sample) and University Hospital, Mangiagalli Hospital, University of Milan (test sample)</p> <p>Place of study: Desio (study sample) and Mangiagalli (test sample), Italy</p> <p>Period of study: July 1998 to December 1999</p> <p>Language: English</p>
Index tests	<p>Index test: TVUS (transvaginal ultrasound)</p> <p>Description of positive case definition by index test as reported: criteria for diagnosis of endometriosis referenced to primary source (Kurjak et al., 1994)</p> <p>Examiners: all pelvic examinations and transvaginal ultrasounds conducted by a single gynaecologist who was not blinded to clinical information and to results of pelvic examination; level of expertise not reported</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: pelvic endometriosis</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 37/90 (41%); rASRM stage I to II 14/37 (38%), rASRM stage III to IV 23/37 (62%)</p>

Eskenazi 2001 (Continued)

Reference standard: laparoscopy 72/90 (80%), laparotomy 18/90 (20%) + histopathology

Description of positive case definition by reference test as reported: diagnosis made by histopathological assessment of biopsied lesions - criteria not provided; surgical procedure described

Examiners: numbers or level of expertise of surgeons not provided; unclear whether blinded to results of index test. All specimens read by a pathologist experienced in histological appearance of endometriosis; slides sent to a second pathologist if visual diagnosis and histological report differed, and if suspected lesion had been biopsied

Flow and timing	Time interval between index test and reference standard: within 34 days Withdrawals: none reported		
Comparative			
Key conclusions by the authors	Non-invasive tools may be used to identify women with ovarian, but not non-ovarian, endometriosis, with excellent agreement with surgical diagnosis		
Conflict of interests	Not reported (supported by grant nos. R82471 from US Environmental Protection Agency, R01 ES07171 and F06 TWO2075-01 from National Institutes of Health, and EA-M1977 from Endometriosis Association)		
Notes	<p>Presented accuracy estimates for transvaginal ultrasound for diagnosis of pelvic endometriosis (only data from prospective arm of the study) confirmed as accurate</p> <p>All cases with positive ultrasound had ovarian involvement; histopathology confirmed 86% of surgically diagnosed endometriosis</p> <p>Predictive algorithm for diagnosis of endometriosis based on history, examination and ultrasound also presented - not included in this review</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	High
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was consid-	Yes		

Eskenazi 2001 (Continued)
ered to be a “positive” re-
sult of index test?

Was the index test per- formed by a single opera- tor or interpreted by con- sensus in a joint session?	Yes	
Were the same clinical da- ta available when the in- dex test results were inter- preted as would be avail- able when the test is used in practice?	Yes	
	Low	Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference stan- dard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference stan- dard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Falco 2011

Study characteristics

Patient sampling	<p>Primary objectives: to evaluate diagnostic accuracy of transvaginal sonography (TVS) for diagnosing deep infiltrating posterior endometriosis (DIPE) and to assess lesion size and test accuracy</p> <p>Study population: patients scheduled for laparoscopy with ≥ 1 symptom suggestive for the presence of endometriosis</p> <p>Selection criteria: not specified</p>
------------------	---

Falco 2011 (Continued)

	Study design: prospective, observational; non-consecutive enrolment
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 65/128, chronic pelvic pain 52/128, infertility 49/128, dyspareunia 41/128, dyschezia 23/128, palpable peritoneal nodules 33/128, ovarian cyst 18/128; previously diagnosed endometriosis 9/128</p> <p>Age: mean 33.6 years, range 18 to 48 years</p> <p>Number enrolled: 128 women</p> <p>Number available for analysis: 96 women</p> <p>Setting: University Hospital "Federico II"</p> <p>Place of study: Naples, Italy</p> <p>Period of study: December 2008 to May 2010</p> <p>Language: Italian</p>
Index tests	<p>Index test: TVUS (TVS)</p> <p>Description of positive case definition by index test as reported: DIPE suspected when irregularly shaped hypoechoic nodules suspicious for endometriotic nodules were located in posterior compartment; described for each site (rectosigmoid, POD, USL, RVS, vagina)</p> <p>Examiners: Operator not unaware of results of bimanual clinical examination but could ask questions about symptoms present; number of operators and level of expertise not provided</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: pelvic endometriosis; DIE (DIPE) - overall and separate anatomical sites</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 76/96 (79.2%), DIPE 52/96 (54.2%)</p> <p>Reference standard: laparoscopy 96/96 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: diagnosis of DIPE defined as presence of endometrial tissue (glands and stroma) on histopathology in at least 1 resected lesion, or direct visualisation of deep endometriotic lesions or obliteration of POD if lesions were unresectable; staging - ASRM classification; surgical procedure described</p> <p>Examiners: numbers or level of expertise of surgeons or pathologists not provided; unclear whether blinded to results of index test</p>
Flow and timing	<p>Time interval between index test and reference standard: within 1 week</p> <p>Withdrawals: 32 (25%) enrolled participants excluded; not explained, presumably did not have surgery</p>
Comparative	
Key conclusions by the authors	<p>In conclusion, this work demonstrates the value of TVS for diagnosis of DIPE and should represent the primary imaging modality in evaluation of patients with suspected endometriosis. Accuracy of this technique depends on knowledge and skill of physician but also on size of endometriotic nodules</p>
Conflict of interests	<p>Not reported</p>
Notes	<p>Reported accuracy estimates for TVS for diagnosis of pelvic endometriosis and DIPE, overall and per each anatomical site, confirmed as accurate</p>

Falco 2011 (Continued)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Falco 2011 (Continued)

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? No

High

Fastrez 2011
Study characteristics

Patient sampling

Primary objectives: to evaluate the value of ^{18}F FDG PET-CT in the diagnosis of endometriosis and to correlate test results with laparoscopic findings

Study population: patients with suspected severe endometriosis (based on clinical presentation) for whom laparoscopy was indicated

Selection criteria: Inclusion criteria: age ≥ 18 years; symptoms consistent with endometriosis, such as chronic pelvic pain and/or dysmenorrhoea resistant to medical therapy and/or infertility. Exclusion criteria: pregnancy or possible pregnancy. All patients had undergone preoperative transvaginal ultrasound and/or MRI

Study design: prospective, observational; non-consecutive enrolment

Patient characteristics and setting

Clinical presentation: dysmenorrhoea 8/10, chronic pelvic pain 1/10, infertility 6/10, dyspareunia 1/10, adnexal mass 1/10; past history of laparoscopy for endometriosis 3/10

Age: mean 31 years, range 21 to 41 years

Number enrolled: 10 women

Number available for analysis: 10 women

Setting: University Hospital CHU St Pierre, Universite Libre de Bruxelles

Place of study: Brussels, Belgium

Period of study: September 2008 to August 2009

Language: English

Index tests

Index test: ^{18}F FDG PET-CT (fluorodeoxyglucose positron emission tomography-computed tomography)

Description of positive case definition by index test as reported: any focal or diffuse ^{18}F FDG uptake above background in location incompatible with normal anatomy and/or physiology considered pathological and correlated with corresponding CT slices. No other specific criteria stated

Examiners: all PET and CT images analysed by the same experienced nuclear medical physicians; number of operators not provided; unclear whether blinded to clinical data and results of other tests

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: pelvic endometriosis

Prevalence of target condition in the sample: pelvic endometriosis 9/10 (90%); rASRM III to IV 6/9 (67%)

Fastrez 2011 (Continued)

Reference standard: laparoscopy 10/10 (100%) + histopathology

Description of positive case definition by reference test as reported:

Peritoneal cavity inspected and lesions of endometriosis described using ASRM classification; histopathological examination ± anti-CD10 immunohistochemistry used to confirm diagnosis. Even in cases in which endometriosis was not confirmed by histology, patients considered to have endometriosis if typical endometriosis lesions observed at visual inspection during laparoscopy. Surgical procedure reported as "classical laparoscopic investigation" but not described

Examiners: Numbers or level of expertise of surgeons or pathologists not provided; laparoscopy performed in blind vs ¹⁸FDG PET-CT data

Flow and timing	Time interval between index test and reference standard: within 1 menstrual cycle Withdrawals: none reported
Comparative	
Key conclusions by the authors	In this preliminary series, we did not observe hypermetabolic activity in relation to endometriosis using ¹⁸ FDG PET-CT. This study's most important limitation is the use of ¹⁸ FDG as an isotopic tracer, which is not specific to endometrial tissue
Conflict of interests	Study authors declared no potential conflicts of interests and received no financial support for research and/or authorship
Notes	Pilot study presenting negative findings; similar results reported by another group in small descriptive study, which did not meet inclusion criteria for this review (Setubal 2011)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	No		

Fastrez 2011 (Continued)

Was the index test performed by a single operator or interpreted by consensus in a joint session?	No
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear
High	
Low	
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Low	
Low	
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Low	

Fedele 1998

Study characteristics	
Patient sampling	<p>Primary objectives: to evaluate the validity of transrectal ultrasonography in assessment of rectovaginal endometriosis</p> <p>Study population: patients scheduled for laparoscopy or laparotomy for pelvic endometriosis, suspected on basis of history and objective findings (not specified)</p> <p>Selection criteria: exclusion criterion: previous surgery for rectovaginal endometriosis</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: infertility 67/140, pelvic pain 52/140; clinical findings 21/140</p> <p>Age: mean 30.2 ± 5.7 years</p>

Fedele 1998 (Continued)

Number enrolled: 140 women

Number available for analysis: 140 women

Setting: University Hospital, The University of Verona

Place of study: Verona, Italy

Period of study: November 1995 to April 1997

Language: English

Index tests	<p>Index test: <u>TRUS</u> (transrectal ultrasonography)</p> <p>Description of positive case definition by index test as reported: presence of deep endometriotic lesions in rectovaginal septum with or without infiltration of rectal or vaginal wall \pm lateral extension to uterosacral ligaments; isolated foci infiltrating uterosacral ligaments not considered diagnostic; sonographic criteria not specified</p> <p>Examiners: ultrasonographer not aware of clinical findings or patient history; knew only that endometriosis was suspected; numbers of examiners and level of expertise not reported</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: posterior DIE (rectovaginal endometriosis) - separate anatomical sites</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 125/140 (89.3%), rectovaginal endometriosis 34/140 (25.3%)</p> <p>Reference standard: laparoscopy 114 (81.4%), laparotomy 26 (18.6%) + histopathology</p> <p>Description of positive case definition by reference test as reported: endometriosis infiltrating rectovaginal septum defined on the basis of surgical and pathological findings, in particular infiltration of the vagina and rectum and lateral infiltration to USL - histological criteria not specified; superficial endometriotic foci on ligaments not considered a diagnostic criterion; surgical procedure described in detail</p> <p>Examiners: numbers or level of expertise of surgeons or pathologists not reported; no blinding to results of index test</p>
Flow and timing	<p>Time interval between index test and reference standard: within 1 week</p> <p>Withdrawals: none reported</p>
Comparative	
Key conclusions by the authors	If our preliminary results are confirmed by larger series, transrectal ultrasonography will be considered a valid diagnostic tool for evaluation of rectovaginal endometriosis
Conflict of interests	Not reported
Notes	<p>Reported accuracy estimates for transrectal ultrasonography for diagnosis of rectovaginal endometriosis and separately for infiltration of vagina, rectum wall and USL confirmed as accurate</p> <p>"The isolated foci on USL were not considered to be a diagnostic criterion", thus may underrepresent both reference standard and index test findings</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Fedele 1998 (Continued)

Was a consecutive or random sample of patients enrolled?	No
Did the study avoid inappropriate exclusions?	Yes
Was a 'two-gate' design avoided?	Yes
<div>HighLow</div>	

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	No
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes
<div>HighLow</div>	

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
<div>HighLow</div>	

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Fedele 1998 (Continued)

Low

Ferrero 2011

Study characteristics

Patient sampling	<p>Primary objectives: to compare the accuracy of multi-detector computerised tomography enteroclysis (MDCT-e) and rectal water contrast transvaginal ultrasonography (RWC-TVS) in determining the presence and extent of bowel endometriosis</p> <p>Study population: patients referred to the endometriosis centre</p> <p>Selection criteria: Inclusion criteria: suspicion of deep pelvic endometriosis (on the basis of gynaecological symptoms and vaginal examination); presence of gastrointestinal symptoms that might be caused by bowel endometriosis; reproductive age; desire to undergo complete surgical excision of the endometriosis. Exclusion criteria: previous bilateral ovariectomy; previous barium radiological examination or other examination for diagnosis of bowel endometriosis; previous bowel surgery (except appendectomy); previous episodes suggestive of intolerance to iodinated contrast medium; renal or hepatic failure; psychiatric disorders</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 72/96, deep dyspareunia 49/96, chronic pelvic pain 61/96, dyschezia 39/96, infertility 32/96, diarrhoea 28/96, constipation 39/96, intestinal cramping 40/96, abdominal bloating 53/96, mucus in the stools 13/96, rectal bleeding 2/96; previous live birth 27/96, previous surgery for endometriosis 39/96, hormonal therapy at time of study 34/96</p> <p>Age: mean 33.4 ± 5.2 years</p> <p>Number enrolled: 96 women</p> <p>Number available for analysis: 96 women</p> <p>Setting: University Hospital: San Martino University Hospital, endometriosis referral centre, Galliera Hospital</p> <p>Place of study: Genoa, Italy</p> <p>Period of study: January 2008 to November 2009</p> <p>Language: English</p>
Index tests	<p>Index test: MDCT-e; TVUS (RWC-TVS)</p> <p>Description of positive case definition by index test as reported: MDCT-e - criterion for diagnosis of bowel endometriosis was the presence of solid nodules with positive enhancement, contiguous or penetrating the thickened intestinal wall or pathological multi-layered appearance of the bowel wall; RWC-TVS - bowel endometriosis appears ultrasonographically as a nodular, solid, hypoechoic lesion, adjacent to and/or penetrating the intestinal wall; unclear whether prespecified criteria or description of findings</p> <p>Examiners: MDCT-e and RWC-TVS were independently and blindly performed by different investigators, who were blinded to the clinical data and knew only that the presence of intestinal endometriosis was suspected; level of expertise not reported</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: bowel endometriosis (ileum - rectum); rectosigmoid endometriosis</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 96/96 (100%); bowel endometriosis 51/96 (53.1%); rectosigmoid endometriosis 48/96 (50%)</p>

Ferrero 2011 (Continued)

Reference standard: laparoscopy 96/96 (100%) + histopathology

Description of positive case definition by reference test as reported: diagnosis and assessment of depth of infiltration of endometriotic nodules referenced to a primary source and described; intestinal endometriosis (ileum - rectum) defined as disease infiltrating at least the muscularis propria; endometriotic foci located on bowel serosa considered peritoneal, not bowel endometriosis; surgical procedure described in detail

Examiners: all surgical procedures performed by a team of gynaecological and colorectal surgeons with extensive experience in the treatment of pelvic and bowel endometriosis, who were aware of index test results. The same pathologist histologically evaluated all specimens excised at surgery; level of expertise not reported

Flow and timing	Time interval between index test and reference standard: within 1 month Withdrawals: none reported
Comparative	
Key conclusions by the authors	Similar accuracy of MDCT-e and RWC-TVS in the diagnosis of rectosigmoid endometriosis, but patients tolerate RWC-TVS better than they do MDCT-e
Conflict of interests	Not reported
Notes	Reported accuracy estimates for MDCT-e and RWC-TVS for diagnosis of overall bowel endometriosis and separately for rectosigmoid confirmed as accurate Accuracy estimates for index tests for diagnosis of various bowel endometriotic lesions, detection rate of lesions of intestinal serosa also presented - not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was consid-	Unclear		

Ferrero 2011 (Continued)

ered to be a “positive” result of index test?

Was the index test performed by a single operator or interpreted by consensus in a joint session?

No

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?

Yes

High

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?

Yes

Were the reference standard results interpreted without knowledge of the results of the index tests?

No

High

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

Yes

Were all patients included in the analysis?

Yes

Low

Ghezzi 2005
Study characteristics

Patient sampling

Primary objectives: to evaluate whether presence of kissing ovaries at ultrasound is a marker for endometriosis, and whether it correlates with severity of the disease

Study population: premenopausal women with adnexal mass or with clinical signs suggestive of pelvic endometriosis who were scheduled for laparoscopic surgery

Ghezzi 2005 (Continued)

Selection criteria: exclusion criteria: previous surgical intervention on adnexa or uterus; history of breast, gastrointestinal tract or genitourinary tract malignancy; history of infertility without symptoms or signs of endometriosis; clinical or ultrasound suspicion of malignancy

Study design: prospective, observational, multi-centre; consecutive enrolment

Patient characteristics and setting

Clinical presentation: chronic pelvic pain, dyspareunia, dysmenorrhoea 309/722, infertility 145/722, adnexal mass not suggestive of endometriosis 413/722

Age: premenopausal, mean age and age range not reported

Number enrolled: 722 women

Number available for analysis: 710 women

Setting: 2 university hospitals: University of Insubria Del Ponte Hospital and University of Berne Hospital

Place of study: Varese, Italy, and Berne, Switzerland

Period of study: January 2000 to November 2003

Language: English

Index tests

Index test: TVUS (transvaginal ultrasound, sign of 'kissing ovaries')

Description of positive case definition by index test as reported: diagnosis of 'kissing ovaries' when both ovaries were joined together behind the uterus in the cul-de-sac and were not separable by pushing the transvaginal probe and by moving the uterus transabdominally; ovarian endometrioma suspected in the presence of round cysts with thick walls, regular margins and homogeneous low echogenicity; presence of definite endometriomas not a prerequisite for the diagnosis of kissing ovaries

Examiners: all ultrasound examinations performed by 3 examiners; level of expertise and blinding to clinical data not reported

Interobserver variability: not provided; each case reviewed by the 3 examiners - in cases of discordant opinion, agreement reached after a collegial discussion of the case

Target condition and reference standard(s)

Target condition: pelvic endometriosis

Prevalence of target condition in the sample: pelvic endometriosis 309/710 (43.5%); rAFS stage I to II 120/309 (39%); rAFS stage III to IV 189/309 (61%)

Reference standard: laparoscopy 710/710 (100%) + histopathology

Description of positive case definition by reference test as reported: visual inspection (rAFS classification) and histological examination; pathological descriptions consistent with endometrial glands and stroma considered endometriosis; descriptions of haemosiderin-laden macrophages alone, although suspected to be endometriosis, not considered clear evidence of the disease; surgical procedure described in detail

Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinded to results of index test

Flow and timing

Time interval between index test and reference standard: within 1 week

Withdrawals: 12 enrolled participants (2%) excluded for the following: uterine myoma misdiagnosed as an adnexal mass (n = 6), malignant ovarian tumour revealed at frozen section examination (n = 4), appendicular mucocoele diagnosed as a sactosalpinx (n = 1), large lymphocyst misdiagnosed as an ovarian cyst (n = 1)

Comparative

Ghezzi 2005 (Continued)

Key conclusions by the authors	Detection of kissing ovaries at ultrasound strongly associated with presence of endometriosis and a marker of the most severe form of this disease
Conflict of interests	Not reported
Notes	<p>Study authors did not estimate the accuracy of 'kissing ovaries' for diagnosis of endometriosis, only association with severity of the disease - data on prediction of severity of endometriosis by index test not included in the review</p> <p>Statement "If isolated peritoneal endometriotic foci were found at surgery in case of a non endometriotic adnexal mass, the patient was classified in accordance with adnexal mass histology" suggests target condition underrepresented by reference standard</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear		
		Unclear	Low

Ghezzi 2005 (Continued)

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
	Unclear
	Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Goncalves 2010
Study characteristics

Patient sampling	<p>Primary objectives: to evaluate use of a specific protocol consisting of transvaginal ultrasonography with bowel preparation (TVUS-BP) to determine number of endometriotic lesions affecting rectosigmoid and depth of these lesions in the bowel wall</p> <p>Study population: patients submitted to laparoscopy on suspicion of endometriosis</p> <p>Selection criteria: inclusion criterion: scheduled to undergo surgery for therapeutic management of endometriosis. Exclusion criterion: any prior bowel surgery</p> <p>Study design: prospective, observational; consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: severe dysmenorrhoea 109/194, deep dyspareunia 120/194, cyclical bowel complaints 112/194, chronic pelvic pain 39/194, infertility 97/194, cyclical urinary complaints 18/194; mean time between onset of symptoms and diagnosis 5.2 years (range 0.4 to 10 years)</p> <p>Age: mean 34.2 ± 4.9 years</p> <p>Number enrolled: 194 women</p> <p>Number available for analysis: 194 women</p> <p>Setting: University Hospital, Sirio Libanes Hospital, University of São Paulo Medical School</p> <p>Place of study: São Paulo, Brazil</p> <p>Period of study: October 2006 to September 2008</p>

Goncalves 2010 (Continued)

Language: English

Index tests	<p>Index test: <u>TVUS</u> (TVUS-BP, with bowel preparation)</p> <p>Description of positive case definition by index test as reported: each examination interpreted in real time and documented in printed photographs; bowel involvement established when a long, nodular, predominantly solid, hypoechogenic lesion adhered to the wall of the intestinal loop; involvement of various layers of the intestinal wall described in detail</p> <p>Examiners: all exams performed by the same radiologist, who was blinded with respect to clinical data and results of other exams to which the patient had been submitted; level of expertise not stated</p> <p>Interobserver variability: not provided</p>		
Target condition and reference standard(s)	<p>Target condition: rectosigmoid endometriosis</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 194/194 (100%): stage I to II 71/194 (37%), stage III to IV 123/194 (63%); rectosigmoid endometriosis 81/194 (42%)</p> <p>Reference standard: laparoscopy 194/194 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: visual inspection (rASRM classification); deep endometriosis of rectum and/or sigmoid confirmed by histology - criteria not specified; surgical procedure described</p> <p>Examiners: surgery performed by the same team in all cases; surgical specimens evaluated by a single pathologist; level of expertise and blinding to results of index test not reported</p>		
Flow and timing	<p>Time interval between index test and reference standard: within 3 months</p> <p>Withdrawals: none reported</p>		
Comparative			
Key conclusions by the authors	TVUS-BP is an adequate exam for evaluating the presence of ≥ 1 rectosigmoid nodule and the deepest layer affected in deep infiltrating bowel endometriosis, confirming the importance of this technique for defining the most appropriate surgical strategy to be implemented		
Conflict of interests	Not reported		
Notes	<p>Reported accuracy estimates for TVUS-BP for diagnosis of rectosigmoid endometriosis confirmed as accurate</p> <p>Diagnostic accuracy in detecting number of lesions and depth of invasion also presented – not included in this review</p> <p>"Women scheduled to undergo surgery for the therapeutic management of endometriosis were included" may imply that a diagnosis of endometriosis was made before enrolment in this study, but the statement is not clear enough for the study to be excluded</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Goncalves 2010 (Continued)

Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Goncalves 2010 (Continued)

Low

Grasso 2010

Study characteristics

Patient sampling	<p>Primary objectives: to compare 2 different imaging modalities - magnetic resonance imaging (MRI) and 3-dimensional sonography (3D-TVUS) - to evaluate their specific role in preoperative work-up of deep infiltrating endometriosis</p> <p>Study population: patients with clinical suspicion of pelvic endometriosis</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: pain (dysmenorrhoea, dyspareunia, chronic pelvic pain) 18/33, infertility 5/33, adnexal masses and/or tenderness at physical examination 10/33</p> <p>Age: mean 35, range 22 to 53 years</p> <p>Number enrolled: 33 women</p> <p>Number available for analysis: MRI 33 women; 3D-TVUS 24 women</p> <p>Setting: University Hospital, Villa Valeria Hospital and Campus Bio Medico University of Rome</p> <p>Place of study: Rome, Italy</p> <p>Period of study: June 2006 to June 2008</p> <p>Language: English</p>
Index tests	<p>Index test: TVUS (3D-TVUS); MRI (T1/T2-w + fat-suppressed + Gd)</p> <p>Description of positive case definition by index test as reported: 3D-TVUS - diagnosis of pelvic endometriosis based on different morphological criteria, which varied for each anatomical location of the disease and included thickening or echogenic nodules or masses with regular or irregular outlines, as described for each site (ovary, USL, posterior vaginal fornix, RVS, sigmoid colon, bladder, POD); MRI - pelvic endometriosis diagnosed when ≥ 1 site of involvement (ovarian or deep pelvic endometriosis) was seen; deeply infiltrating pelvic endometriosis defined by the presence of endometriosis in 1 of the following areas: torus uterinus and USL, vagina, rectovaginal septum, sigmoid colon, ureters and bladder - criteria described for each site</p> <p>Examiners: all 3D-TVUS scans performed by a gynaecologist with 20 years' experience with endometriosis and gynaecological ultrasound, who was blinded to the patient's clinical history, symptoms and MR results; MRI analysed prospectively by 1 radiologist who was blinded to clinical and sonographic findings; level of expertise not reported.</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: DIE (deeply infiltrating pelvic endometriosis)</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 33/33 (100%), deeply infiltrated pelvic endometriosis 26/33 (78.7%)</p> <p>Reference standard: laparoscopy 33/33 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: deeply infiltrating endometriosis defined on the basis of surgical and/or pathological findings as follows: posterior compartment alone: USL, vagina, bowel, sigmoid colon, rectovaginal septum, obliteration of the pouch of Douglas,</p>

Imaging modalities for the non-invasive diagnosis of endometriosis (Review)

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

138

Grasso 2010 (Continued)

ureters; anterior compartment alone: bladder; both anterior and posterior parts of compartment; histological criteria not specified; surgical procedure described

Examiners: numbers or level of expertise of surgeons not provided; 2 different pathologists analysed the first 20 patients and the last 13 patients, respectively; level of expertise not stated; unclear whether blinded to results of the index test

Flow and timing	Time interval between index test and reference standard: 1 to 4 weeks Withdrawals: 9 enrolled participants excluded from ultrasound group - referred to other institutions for ultrasound test		
Comparative			
Key conclusions by the authors	MR accurately diagnoses deep infiltrating endometriosis; 3D-US accurately diagnoses deep infiltrating endometriosis in specific locations		
Conflict of interests	Not reported		
Notes	<p>Reported summary estimates for 3D-TVS for diagnosis of deep pelvic endometriosis not confirmed; reported accuracy estimates for MRI for diagnosis of deep pelvic endometriosis confirmed as accurate</p> <p>Accuracy estimates also presented separately for specific sites of deep endometriosis (USL, posterior vaginal fornix, rectovaginal septum, sigmoid colon and bladder) and for endometriomas, which were calculated per number of lesions - not included in this review as calculated per number of lesions rather than number of patients</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a “positive” result of index test?	Yes		

Grasso 2010 *(Continued)*

Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes	
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes	
	Low	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Guerriero 1996a
Study characteristics

Patient sampling	<p>Primary objectives: to evaluate the accuracy of transvaginal sonography used in combination with Ca19.9 assay (with or without Ca-125 determinations) for the diagnosis of endometrioma in premenopausal women</p> <p>Study population: women scheduled for laparoscopy or laparotomy for a persistent ovarian mass</p> <p>Selection criteria: Inclusion criteria: premenopausal, non-pregnant women</p> <p>Study design: prospective, observational; consecutive enrolment</p>
------------------	--

Guerriero 1996a (Continued)

Patient characteristics and setting	<p>Clinical presentation: symptoms and clinical findings: persistent adnexal mass 118/118 (100%), infertility 45/118 (53%)</p> <p>Age: mean 33.3 ± 9.6 years, range 14 to 54 years</p> <p>Number enrolled: 118 women</p> <p>Number available for analysis: 118 women</p> <p>Setting: University Hospital, University of Cagliari</p> <p>Place of study: Cagliari, Italy</p> <p>Period of study: November 1994 to November 1995</p> <p>Language: English</p>		
Index tests	<p>Index test: <u>TVUS</u> (transvaginal ultrasonography)</p> <p>Description of positive case definition by index test as reported: criteria for diagnosis of endometrioma as previously described (referenced to Mais et al., 1993; Kurjak et al., 1994); all examinations performed in the follicular phase of the cycle</p> <p>Examiners: all scans performed by the same physician; level of expertise and blinding to clinical data not reported</p> <p>Interobserver variability: not provided</p>		
Target condition and reference standard(s)	<p>Target condition: ovarian endometriosis</p> <p>Prevalence of target condition in the sample: ovarian endometriosis 39/118 (33%)</p> <p>Reference standard: laparoscopy 99/118 (84%), laparotomy 19/118 (16%) + histopathology</p> <p>Description of positive case definition by reference test as reported: visual inspection + histological confirmation - criteria not stated; surgical procedure described</p> <p>Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinded to results of the index test</p>		
Flow and timing	<p>Time interval between index test and reference standard: within 2 days</p> <p>Withdrawals: none reported</p>		
Comparative			
Key conclusions by the authors	Transvaginal ultrasonography used alone is the most cost-effective method in the preoperative differential diagnosis of endometrioma		
Conflict of interests	Not reported		
Notes	<p>Reported accuracy estimates for transvaginal sonography for detection of endometrioma confirmed as accurate (based on number of patients)</p> <p>Accuracy estimates for a combination of transvaginal ultrasonography with Ca-19.9 and Ca-125 also presented – not included in this review</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns

Guerriero 1996a (Continued)

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?	Yes	
Did the study avoid inappropriate exclusions?	Yes	
Was a 'two-gate' design avoided?	Yes	
	Low	Low

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes	
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes	
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear	
	Unclear	Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Guerriero 1996b
Study characteristics

Patient sampling	<p>Primary objectives: to assess the role of transvaginal ultrasonography in combination with Ca-125 plasma levels in diagnosis of endometrioma</p> <p>Study population: women who were submitted to laparoscopy or laparotomy because of the presence of a persistent adnexal mass</p> <p>Selection criteria: Inclusion criteria: premenopausal, non-pregnant women</p> <p>Study design: prospective, observational; consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: not specified</p> <p>Age: range 20 to 49 years, mean not provided</p> <p>Number enrolled: 101 women</p> <p>Number available for analysis: 101 women</p> <p>Setting: University Hospital, University of Cagliari</p> <p>Place of study: Cagliari, Italy</p> <p>Period of study: November 1993 to October 1994</p> <p>Language: English</p>
Index tests	<p>Index test: <u>TVUS</u> (transvaginal ultrasonography)</p> <p>Description of positive case definition by index test as reported: endometrioma defined in accordance with previously published criteria (referenced to Mais et al., 1993) and described; all examinations performed in the follicular phase of the cycle</p> <p>Examiners: all scans performed by the same physician; level of expertise and blinding to clinical data not reported</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: ovarian endometriosis</p> <p>Prevalence of target condition in the sample: ovarian endometriosis 29/101 (28.7%)</p> <p>Reference standard: laparoscopy, laparotomy (number for each group not reported) + histopathology</p> <p>Description of positive case definition by reference test as reported: ovarian masses identified as endometriomas on histopathology when ≥ 2 of the following findings were present: endometrial surface epithelium, endometrial glands or gland-like structures, endometrial stroma and haemosiderin-laden macrophages; surgical procedure described</p> <p>Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinded to results of the index test</p>
Flow and timing	<p>Time interval between index test and reference standard: within 2 days</p> <p>Withdrawals: none reported</p>
Comparative	

Guerriero 1996b (Continued)

Key conclusions by the authors	Transvaginal ultrasonography used alone has better predictive capacity than combined methods for differentiating endometriomas from other adnexal masses
Conflict of interests	Not reported
Notes	Reported accuracy estimates for transvaginal sonography for diagnosis of ovarian endometriosis confirmed as accurate (based on number of patients)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			

Guerriero 1996b (Continued)

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Guerriero 2007
Study characteristics

Patient sampling	<p>Primary objectives: to determine the accuracy of transvaginal ultrasonography (TVUS) by using this modified and 'tenderness-guided' approach in the diagnosis of deep endometriosis of the cul-de-sac, retrocervical region and rectovaginal septum</p> <p>Study population: women scheduled for laparoscopic surgery for rectovaginal endometriosis, suspected on the basis of patient history of pelvic pain and/or clinical examination</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational; consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: pelvic pain in all 50 women: dyspareunia 19/50, dysmenorrhoea 42/50, infertility 5/50; previous medical treatment for persistent pelvic pain (estrogens, progestins and/or gonadotropin-releasing hormone agonist and non-steroidal anti-inflammatory drugs) for ≥ 2 years 50/50</p> <p>Age: mean 33 ± 5 years, range 22 to 41 years</p> <p>Number enrolled: 50 women</p> <p>Number available for analysis: 50 women</p> <p>Setting: University Hospital, University of Cagliari</p> <p>Place of study: Cagliari, Italy</p> <p>Period of study: January 2005 to May 2005</p> <p>Language: English</p>
Index tests	<p>Index test: <u>TVUS</u> (TVUS tenderness-guided approach)</p> <p>Description of positive case definition by index test as reported: diagnostic criteria (referenced to Bazot et al., 2003; Guerriero et al., 1998) described; assessed areas: rectouterine pouch, rectovaginal septum and POD; suspected involvement of intestine and partial or complete obliteration of POD recorded. On the basis of ultrasonographic images, rectovaginal endometriosis also graded using the scoring system of Adamyan and described</p> <p>Examiners: all scans performed by 1 investigator, who has had more than 15 years of experience with TVUS; unclear whether blinded to clinical data</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: posterior DIE (deep posterior endometriosis); ovarian endometriosis</p>

Guerriero 2007 (Continued)

Prevalence of target condition in the sample: pelvic endometriosis 43/50 (86%); deep posterior endometriosis 31/50 (62%); ovarian endometrioma 9, unclear whether numbers of lesions or patients

Reference standard: laparoscopy 50/50 (100%) + histopathology

Description of positive case definition by reference test as reported: visual inspection ± histopathological examination, criteria referenced to primary source and described; surgical procedure not described

Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinded to results of the index test

Flow and timing	Time interval between index test and reference standard: within 7 days		
	Withdrawals: none reported		
Comparative			
Key conclusions by the authors	Our new TVUS approach appears to be an accurate, inexpensive and less invasive method for the diagnosis of deep endometriosis		
Conflict of interests	Not reported		
Notes	<p>Reported accuracy estimates for transvaginal tenderness-guided ultrasonography for diagnosis of posterior deep endometriosis confirmed as accurate</p> <p>No data available for calculating test accuracy for separate diagnosis of vaginal and rectal wall involvement</p> <p>Unclear whether accuracy estimates for diagnosis of endometrioma are based on numbers of cysts or numbers of patients; given 100% accuracy, unlikely the way of calculation matters; therefore these data are presented</p> <p>Accuracy estimates for the index test for staging of endometriosis and concordance with surgical staging also reported - not presented in this review</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Guerriero 2007 *(Continued)*

Did the study provide a clear pre-specified definition of what was considered to be a “positive” result of index test?	Yes	
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes	
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear	
		Unclear Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
		Unclear Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
		Low

Guerriero 2008
Study characteristics

Patient sampling	<p>Primary objectives: to evaluate the diagnostic accuracy of transvaginal tenderness-guided ultrasonography (tg-TVUS) in identification of the location of deep endometriotic implants</p> <p>Study population: women scheduled for laparoscopic surgery for clinically suspected endometriosis on the basis of patient history of pelvic pain and/or clinical examination</p> <p>Selection criteria: not specified</p>
------------------	---

Guerriero 2008 (Continued)

	Study design: prospective, observational; consecutive enrolment
Patient characteristics and setting	<p>Clinical presentation: pelvic pain in all 88 patients: dyspareunia 40/88, dysmenorrhoea 71/88, infertility 10/88; previous medical treatment for persistent pelvic pain (estrogens, progestins and/or GnRH agonist and non-steroidal anti-inflammatory drugs) for ≥ 2 years 88/88</p> <p>Age: mean 33 ± 5 years, range 20 to 45 years</p> <p>Number enrolled: 88 women</p> <p>Number available for analysis: 88 women</p> <p>Setting: University Hospital, University of Cagliari</p> <p>Place of study: Cagliari, Italy</p> <p>Period of study: December 2005 to December 2007</p> <p>Language: English</p>
Index tests	<p>Index test: TVUS (tg-TVUS)</p> <p>Description of positive case definition by index test as reported: deep endometriosis implants suspected from the presence of hypoechoic linear thickening or nodules/masses with or without regular contours in 5 locations: vaginal walls, RVS, rectosigmoid involvement, USL and anterior compartment (anterior pouch and/or bladder)</p> <p>Examiners: all scans performed by 1 investigator who had more than 15 years' experience with transvaginal ultrasonography at the outset of the study; unclear whether blinded to clinical data</p> <p>Interobserver variability: reproducibility of the technique determined by evaluation of 10 symptomatic patients by 2 examiners, each with a different level of expertise in ultrasonography in gynaecology; intraobserver agreement good or very good for both examiners with different degrees of experience (kappa values ranging from 0.70 to 0.88)</p>
Target condition and reference standard(s)	<p>Target condition: DIE (deep pelvic endometriosis) - separate anatomical sites</p> <p>Prevalence of target condition in the sample: deep pelvic endometriosis 72/88 (81.8%)</p> <p>Reference standard: laparoscopy 88/88 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: visual inspection \pm histopathological examination as previously reported (referenced to Bazot et al., 2004) and described; surgical procedure not described</p> <p>Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinding to results of the index test</p>
Flow and timing	<p>Time interval between index test and reference standard: within 1 week</p> <p>Withdrawals: none reported</p>
Comparative	
Key conclusions by the authors	<p>This technique shows high specificity and sensitivity in the detection of vaginal and rectovaginal endometriosis. Good specificity associated with lower sensitivity obtained in the diagnosis of deep endometriosis of uterosacral ligaments, rectosigmoid involvement or anterior deep endometriosis</p>
Conflict of interests	<p>Study authors declared no conflict of interest</p>
Notes	<p>Reported accuracy estimates for transvaginal tenderness-guided ultrasonography for diagnosis of specific sites of deep pelvic endometriosis confirmed as accurate</p>

Guerriero 2008 (Continued)

Accuracy estimates for bladder endometriosis reported by study authors but not presented in the review, because this was not an assessed target condition

No data available for calculating test accuracy for overall deep pelvic endometriosis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Guerriero 2008 (Continued)

	Unclear	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Guerriero 2014

Study characteristics	
Patient sampling	<p>Primary objectives: to compare diagnostic performance of 2D and 3D ultrasonography (2D-US vs 3D-US) in detecting DIE in women with clinical suspicion and to assess the reproducibility of 3D-US</p> <p>Study population: all premenopausal women with clinical suspicion of deep endometriosis who were scheduled for surgery in our department</p> <p>Selection criteria: Inclusion criteria: reproductive age, clinically suspected endometriosis; exclusion criteria: abdominal mass larger than 10 cm with distortion of pelvic anatomy, emergency laparoscopy due to acute pain, 2D-US or 3D-US not performed, insufficient description at surgery, pregnancy at time of diagnosis, surgery longer than 30 days after ultrasound</p> <p>Study design: prospective, observational, diagnostic; consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: chronic pelvic pain 101/202, dyspareunia 51/202, dysmenorrhoea 132/202; previous surgery for pelvic pain 20/202; hormonal treatment at the time of ultrasound examination 43/202</p> <p>Age: mean 34 ± 6 years, range 18 to 52 years</p> <p>Number enrolled: 240 women</p> <p>Number available for analysis: 202 women</p> <p>Setting: University Hospital, Ospedale San Giovanni di Dio, University of Cagliari</p> <p>Place of study: Cagliari, Italy</p> <p>Period of study: January 2009 to September 2012</p> <p>Language: English</p>
Index tests	<p>Index test: <u>TVUS 2 types</u> (2D-US (tg-TVUS) and 3D-US)</p> <p>Description of positive case definition by index test as reported: <u>2D-US</u> - described separately for each anatomical location (referenced to Guerriero et al., 1998, 2007, 2008; Bazot et al., 2004a, b; Abrao et al., 2007; Hudelist et al., 2011a,b); <u>3D-US</u> - vaginal and rectovaginal endometriosis appearing as small irregular nodules (evaluated using a sagittal plane); rectosigmoid lesions appearing as spiculated lesions with retracting lines all around the nodule; uterosacral lesions showing a nodular plaque shape laterally to the uterine torus (evaluated in coronal plane)</p>

Guerriero 2014 (Continued)

Examiners: all scans performed by 1 investigator who had more than 20 years' experience with transvaginal ultrasonography. Unclear whether operator was blinded to clinical data

Interobserver variability: performed for 3D-US only on random sample of images from 35 patients by 2 operators (experienced and less experienced) who were blinded to clinical data and previous results. Interobserver agreement 0.7094 (kappa analysis); intraobserver agreement good or very good for both examiners with different degrees of experience (kappa values ranging from 0.8754 for expert reader to 0.7087 for less experienced reader)

Target condition and reference standard(s)

Target condition: posterior DIE (deep pelvic endometriosis) - separate anatomical sites (rectosigmoid and other posterior, including USL, vaginal fornices, RVS)

Prevalence of target condition in the sample: deep pelvic endometriosis 129/202 (64%) participants: single nodule 75/129 (58%), ≥ 1 location endometriosis 54/129 (42%); posterior DIE 122/129 (95%), rectosigmoid endometriosis 77/129 (60%), complete obliteration of POD 51/129 (40%)

Reference standard: laparoscopy 194/202 (96%), laparotomy 8/202 (4%) + histopathology

Description of positive case definition by reference test as reported: visual inspection \pm histopathological examination as previously reported (referenced to a primary source and described); surgical procedure not described

Examiners: same group of surgeons with ≥ 10 years' experience. Not reported whether surgeons blinded to imaging results. Numbers or level of expertise of pathologists not reported

Flow and timing

Time interval between index test and reference standard: within 1 month

Withdrawals: 38 (16%) patients were not included in the study: 3 had ovarian mass > 10 cm; 3 had undergone emergency laparoscopy; 24 underwent only 2DUS; for 7, description of surgery was insufficient; for 1, surgery was performed longer than 1 month post ultrasound

Comparative

Key conclusions by the authors

Study shows that 3D-US is a useful new technique not only for preoperative evaluation of DIE, but also for follow-up of expectant management or medical treatment

Conflict of interests

Study authors declared no conflict of interest. Funding: partially supported by the Regione Autonoma della Sardegna (project code CPR-24750)

Notes

Reported accuracy estimates correct for transvaginal 2D ultrasonography and 3D ultrasonography for diagnosis of posterior deep pelvic endometriosis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	Low

Guerriero 2014 (Continued)

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test? Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice? Yes

Low
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? No

High
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? No

High

Ha 1994

Study characteristics

Patient sampling	<p>Primary objectives: to compare fat-suppressed T1-w MRI with conventional MR images for diagnosis of endometriosis, focusing on detectability of peritoneal implants</p> <p>Study population: patients with suspected endometriosis</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: not specified</p> <p>Age: mean 35 years, range 20 to 52 years</p> <p>Number enrolled: 31 women</p> <p>Number available for analysis: 31 women</p> <p>Setting: University Hospital, Catholic University Medical College</p> <p>Place of study: Seoul, Korea</p> <p>Period of study: 12-month period, dates not specified</p> <p>Language: English</p>
Index tests	<p>Index test: <u>MRI 2 types</u> (T1/T2-w MRI; fat-suppressed T1-w MRI)</p> <p>Description of positive case definition by index test as reported: diagnostic criteria as previously published (referenced to Arrive et al., 1989; Togashi et al., 1991; Nishimura et al., 1987; Zawin et al., 1989) and described; pelvic adhesions excluded from analysis because fat-suppressed images were not useful in detecting fibrotic lesions; conventional and fat-suppressed images evaluated separately in random order</p> <p>Examiners: images reviewed independently by 2 radiologists; level of expertise not reported. Observer knew only that patients had suspected endometriosis</p> <p>Interobserver variability: not provided; consensus findings used if interpretations differed</p>
Target condition and reference standard(s)	<p>Target condition: pelvic endometriosis</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 29/31 (94%); rASRM stage I 7/29 (24%), stage II 7/29 (24%), stage III 4/29 (14%), stage IV 11/29 (38%)</p> <p>Reference standard: laparoscopy 31/31 (100%)</p> <p>Description of positive case definition by reference test as reported: endometriosis diagnosed on the basis of visualisation of pelvic cavity at laparoscopy as endometrial cysts or peritoneal implants (rASRM classification); anatomical sites of involvement divided into 6 categories: right or left ovary, right or left uterine surface or uterosacral ligament, cul-de sac and other anatomic sites (most often, rectum)</p> <p>Examiners: numbers or level of expertise of surgeons not reported; unclear whether blinded to results of the index test</p>
Flow and timing	<p>Time interval between index test and reference standard: within 2 weeks</p> <p>Withdrawals: none reported</p>
Comparative	

Ha 1994 (Continued)

Key conclusions by the authors	Results show fat-suppressed MR imaging as more accurate in the diagnosis of pelvic endometriosis and better than conventional MR imaging for predicting severity of disease
Conflict of interests	Not reported
Notes	<p>Reported accuracy estimates for MRI, both conventional and fat-suppressed, for diagnosis of pelvic endometriosis not confirmed</p> <p>Accuracy estimates for specific sites of endometriosis as well as for peritoneal and ovarian disease not included in this review as calculated per number of lesions rather than number of patients</p> <p>Data on accuracy of MRI in predicting severity of disease also presented – not included in this review</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			

Ha 1994 (Continued)

Is the reference standards likely to correctly classify the target condition?	Unclear
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Unclear	
Low	
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Low	

Holland 2010
Study characteristics

Patient sampling	<p>Primary objectives: to examine the ability of preoperative transvaginal ultrasound (TVUS) scanning to assess the severity of pelvic endometriosis</p> <p>Study population: women with clinically suspected or proven pelvic endometriosis</p> <p>Selection criteria: inclusion criteria: premenopausal women with clinical suspicion of endometriosis awaiting diagnostic laparoscopy; women diagnosed with pelvic endometriosis at diagnostic laparoscopy awaiting operative treatment; age ≥ 16 years; ability to provide informed consent. Exclusion criteria: women who could not undergo TVUS scan; women who became pregnant whilst awaiting surgery</p> <p>Study design: observational, multi-centre; prospective consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 142/201, chronic pelvic pain 104/201, dyspareunia 78/201, infertility 38/201, dyschezia 7/201, cyclical rectal bleeding 2/201; single presenting symptom present in 72/201, 2 presenting symptoms in 78/201 and ≥ 3 symptoms in 51/201</p> <p>Age: mean 34.9 ± 6.79 years (95% CI 33.98 to 35.86), range 19 to 51 years</p> <p>Number enrolled: 211 women</p> <p>Number available for analysis: 201 women</p> <p>Setting: University Hospital, King's College Hospital</p> <p>Place of study: London, UK</p> <p>Period of study: July 2006 to December 2008</p> <p>Language: English</p>

Holland 2010 (Continued)

Index tests	<p>Index test: TVUS (TVS)</p> <p>Description of positive case definition by index test as reported: for ovarian endometrioma, DIE and rectosigmoid endometriosis diagnostic criteria (referenced to a primary source) described. Adhesions defined as minimal when ovary could be mobilised from most (> 2/3) of the surrounding structures, moderate when ovarian mobility was reduced but structures on 2/3 to 1/3 of the surface of the ovary and severe when fixed ovaries could not be mobilised at all or separated from surrounding structures; complete obliteration of POD assessed as the absence of sliding between the serosa on the posterior surface of the cervix or uterus and the bowel; partial obliteration present when some free sliding was seen, or when adnexal structures were firmly adherent to the posterior aspect of the uterus but the bowel appeared to be free</p> <p>Examiners: TVS examination performed by 4 ultrasound operators who were all gynaecologists with a high level of expertise in gynaecological ultrasonography. Ultrasound operators blinded to previous surgical findings. Examiner A performed 104 (51.7%), examiner B performed 68 (33.8%), examiner C performed 18 (9%) and examiner D performed 11 (5.5%) examinations</p> <p>Interobserver variability: accuracy estimates for diagnosis of pelvic endometriosis calculated separately for the 2 examiners and compared - no significant difference found in overall accuracy between these 2 examiners</p>
Target condition and reference standard(s)	<p>Target condition: pelvic endometriosis; DIE - overall and separately for anterior and posterior compartments; POD obliteration</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 139/201 (69.2%); DIE 71/201 (35.3%)</p> <p>Reference standard: laparoscopy 201/201 (100%)</p> <p>Description of positive case definition by reference test as reported: endometriotic lesions identified by visualisation of pelvic cavity (ASRM classification); surgical procedure described</p> <p>Examiners: all patients operated on by 4 different laparoscopic surgeons with a high level of expertise in laparoscopic surgery; surgeons were blinded to detailed TVS findings</p>
Flow and timing	<p>Time interval between index test and reference standard: mean interval 37.5 ± 23.2 days (95% CI 34.3 to 40.8; SD 23.2), range 0 to 87 days</p> <p>Withdrawals: 10 (5%) enrolled participants excluded for the following: 5 became pregnant whilst awaiting surgery, 1 cancelled her operation, 1 underwent unsuccessful laparoscopy and 3 were lost to follow-up</p>
Comparative	
Key conclusions by the authors	TVS is a good test for assessing the severity of pelvic endometriosis. TVS is particularly accurate in detecting severe disease, which could facilitate effective triaging of women for appropriate surgical care
Conflict of interests	Not reported
Notes	<p>Reported accuracy estimates for TVS for diagnosis of pelvic endometriosis and for DIE overall or separate for each site confirmed as accurate</p> <p>Data on TVS staging of endometriosis and on correlation of ultrasound and laparoscopic assessment of the severity of pelvic endometriosis also reported - not presented in this review</p> <p>In addition, study authors compared performance of Examiners A and B in diagnosing severe pelvic endometriosis using ultrasound - not presented in this review</p> <p>'Women diagnosed with pelvic endometriosis at diagnostic laparoscopy awaiting operative treatment' as one of the inclusion criterion may imply that some participants were diagnosed before enrolment in the study; no data on previous surgery for endometriosis; size of this subgroup unclear</p>
Methodological quality	

Holland 2010 (Continued)

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	No		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowl-	Yes		

Holland 2010 (Continued)

edge of the results of
the index tests?

	Low	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Hottat 2009
Study characteristics

Patient sampling	<p>Primary objectives: to determine the accuracy of 3.0T pelvic magnetic resonance (MR) imaging in preoperative assessment of endometriosis; to evaluate colon wall involvement after intrarectal gel administration</p> <p>Study population: patients referred for pelvic MR imaging because of clinical suspicion of endometriosis</p> <p>Selection criteria: exclusion criteria: common contraindications to MRI (pacemaker, metallic foreign bodies, claustrophobia), age < 18 years, postmenopausal status</p> <p>Study design: observational; prospective consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 19/41, chronic pelvic pain 29/41, dyspareunia 5/41, suspicious clinical examination 15/41, past hx of endometriosis 7/41</p> <p>Age: mean 33 years, range 20 to 46 years</p> <p>Number enrolled: 106 women</p> <p>Number available for analysis: 41 women</p> <p>Setting: endometriosis referral centre, Erasme Hospital, Université Libre de Bruxelles</p> <p>Place of study: Brussels, Belgium</p> <p>Period of study: March 2007 to August 2008</p> <p>Language: English</p>
Index tests	<p>Index test: <u>MRI</u> (3.0T Magnetom system (3.0T MRI))</p> <p>Description of positive case definition by index test as reported: systematic analysis of the pelvic cavity performed, and locations of lesions determined; investigated locations included uterus, adnexa, POD, USLs, vagina, small bowel, colon wall, vesicouterine pouch, bladder and</p>

Hottat 2009 (Continued)

ureters; DIE described as nodular or retractor-like fibrotic-like hypointense tissue on T2-w and isointense to muscle on T1-w images; endometrioma described as a cystic adnexal lesion with hyperintensity on fat-suppressed T1-w and 'shading' on T2-w images; for colon involvement, precise location and infiltration described

Examiners: 2 investigators with 8 years' and 1 year experience in MRI; blinded to clinical findings; independently and prospectively analysed all images

Interobserver variability: level of agreement between the 2 readers reported for each site of endometriosis

Target condition and reference standard(s)	<p>Target condition: pelvic endometriosis; DIE - overall and separately for specific anatomical locations</p> <p>Prevalence of target condition in the sample: DIE 27/41 (66%); USL 21/41, POD 22/41, vaginal 11/41, colon 13/41</p> <p>Reference standard: laparoscopy 34/41; laparotomy 7/41 + histopathology (100%)</p> <p>Description of positive case definition by reference test as reported: endometriotic lesions identified by visualisation of pelvic cavity with subsequent histological confirmation; surgical procedure described</p> <p>Examiners: both surgeon and pathologist with more than 10 years' experience in evaluation of endometriosis; same team for all cases</p>
Flow and timing	<p>Time interval between index test and reference standard: mean interval 60 days, range 2 to 105 days</p> <p>Withdrawals: 65 (61%) enrolled participants excluded as they did not undergo surgery</p>
Comparative	
Key conclusions by the authors	MR imaging of the pelvis at 3.0T is accurate in the diagnosis and staging of deep endometriosis for preoperative assessment of patients clinically suspected of having endometriosis
Conflict of interests	Study authors stated no financial relationship to disclose
Notes	<p>Reported accuracy estimates for MRI for diagnosis of DIE overall or separate for each site (ovarian, USL, vaginal, RS, anterior DIE, POD) confirmed as accurate</p> <p>Data for ovarian and USL endometriosis reported per patient</p> <p>In addition, study authors compared performance of Examiners 1 and 2 - not presented in this review (only data for experience examiner reported)</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		

Hottat 2009 (Continued)

	Low	Low
DOMAIN 2: Index Test Any test		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes	
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes	
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes	
	Low	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes	
	Low	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
	High	

Hudelist 2011a

Study characteristics

Patient sampling	<p>Primary objectives: to compare the diagnostic performance of clinical vaginal examination vs transvaginal sonography (TVS) in the presurgical diagnosis of DIE</p> <p>Study population: women with suspected endometriosis attending 1 of 3 pelvic pain clinics who were referred to the pelvic pain clinic for laparoscopy because of suspected endometriosis on the basis of clinical history and the referring physician's clinical findings, or were self referred (coming to the pain clinic without seeing any gynaecologist before this time for their current problems)</p> <p>Selection criteria: inclusion criterion: premenopausal women; exclusion criteria: history of gynaecological cancer; previous surgery for deep infiltrating endometriosis or other disease entities requiring resection of the bladder and/or dissection of the rectovaginal space and/or anterior rectosigmoidal wall; inability to perform TVS (congenital abnormalities of the genital tract or virginity)</p> <p>Study design: prospective, observational, multi-centre; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 111/129, dyspareunia 72/129, dyschezia 39/129, dysuria 6/129, chronic pelvic pain 45/129, subfertility 20/129</p> <p>Age: mean 32.2 ± 5.4 years, range 17 to 44 years</p> <p>Number enrolled: 153 women</p> <p>Number available for analysis: 129 women</p> <p>Setting: 3 tertiary referral service Hospitals: Worthing and Southlands Hospital, Ashford and St Peters Hospital, Villach Hospital (endometriosis centre)</p> <p>Place of study: Villach, Austria; Worthing and Chertsey, UK</p> <p>Period of study: not stated</p> <p>Language: English</p>
Index tests	<p>Index test: TVUS (TVS)</p> <p>Description of positive case definition by index test as reported: diagnosis of endometrioma based on the presence of a cyst or multiple cysts containing diffuse low-level echoes; diagnosis of pelvic endometriosis based on different morphological criteria that varied for each anatomical location of the disease and included regular or irregular hypoechoic nodular structure, cystic mass or hypoechoic linear thickening with regular or irregular margins, described for each site (USL, vaginal wall, RVS, bladder, rectosigmoid colon); POD obliteration considered complete when uterus, adnexa and rectosigmoid colon were adherent, with disappearance of the peritoneal structure, and incomplete when peritoneal limits were partially identified by the presence or absence of suspended or lateralised fluid collection</p> <p>Examiners: all TVS scans performed by 1 experienced examiner who was blinded to results of the vaginal examinations but was aware that women were being investigated for chronic pelvic pain; therefore, endometriosis was suspected</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: DIE - separate anatomical sites; ovarian endometriosis</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 83/129 (64.3%); DIE 52/129 (40.3%); ovarian endometriosis 27/129 (16.2%)</p> <p>Reference standard: laparoscopy 129/129 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: DIE defined as subperitoneal endometriotic infiltration of tissues > 5 mm; histological presence of endometriosis taken to represent a 'true positive' diagnosis of endometriosis - histological criteria not specified; surgical procedure described in detail</p>

Hudelist 2011a (Continued)

Examiners: total of 3 surgeons performed laparoscopy, all of whom had more than 10 years' experience in radical laparoscopic surgery for DIE and were blinded to results of the vaginal examination and TVS at 1 of the centres but were aware of the vaginal examination and TVS results at the other 2 centres; numbers and level of expertise of pathologists not reported

Flow and timing

Time interval between index test and reference standard: within 3 months (personal communication with study author)

Withdrawals: 24 patients excluded because they did not meet the inclusion criteria: 18 had a history of previous surgery for DIE, 3 had a history of gynaecological cancer and 3 were virgins

Comparative

Key conclusions by the authors

TVS is a more useful test than vaginal examination for detecting endometriosis in the ovaries and rectosigmoid, and can be recommended as the method of choice for primary and preoperative assessment of pelvic pain patients with suspected endometriosis

Conflict of interests

Not reported

Notes

Reported accuracy estimates for TVS for diagnosis of specific sites of DIE and endometrioma confirmed as accurate

Data for overall DIE irrespective of the site not available

Accuracy estimates for bladder endometriosis reported by study authors but not presented in the review because this was not an assessed target condition

Diagnostic performance of vaginal examination and combination of pelvic examination with TVS for preoperative diagnosis of endometriosis also reported – this was not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was	Yes		

Hudelist 2011a *(Continued)*

considered to be a
“positive” result of in-
dex test?

Was the index test per-
formed by a single op-
erator or interpreted
by consensus in a joint
session?

Yes

Were the same clinical
data available when
the index test results
were interpreted as
would be available
when the test is used
in practice?

Yes

Low

Low

DOMAIN 3: Reference Standard

Is the reference stan-
dards likely to correct-
ly classify the target
condition?

Yes

Were the reference
standard results inter-
preted without knowl-
edge of the results of
the index tests?

No

High

Low

DOMAIN 4: Flow and Timing

Was there an appro-
priate interval be-
tween index test and
reference standard?

Yes

Did all patients receive
the same reference
standard?

Yes

Were all patients in-
cluded in the analysis?

Yes

Low

Hudelist 2013
Study characteristics

Patient sampling

Primary objectives: to describe a simple diagnostic sign that could be used to triage women with mild vs advanced endometriosis affecting the rectosigmoid

Hudelist 2013 (Continued)

	<p>Study population: women attending pelvic pain clinic with suspected endometriosis and scheduled for laparoscopy on the basis of clinical examination and TVS findings</p> <p>Selection criteria: exclusion criteria: history of gynaecological cancer; previous surgery for deep infiltrating endometriosis or other disease entities requiring resection of the bladder and/or dissection of the rectovaginal space and/or anterior rectosigmoidal wall; inability to perform TVS (congenital abnormalities of the genital tract or virginity) or non-availability of consent</p> <p>Study design: prospective, observational, multi-centre; consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 116/117, dyspareunia 74/117, dyschezia 31/117, dysuria 9/117, chronic pelvic pain 32/117, subfertility 22/117</p> <p>Age: mean 31.6 ± 6.5 years</p> <p>Number enrolled: 142 women</p> <p>Number available for analysis: 117 women</p> <p>Setting: Department of O&G, Stage III Center for Endometriosis & Pelvic Pain, Wilhelminen Hospital</p> <p>Place of study: Vienna, Austria</p> <p>Period of study: July 2011 to May 2012</p> <p>Language: English</p>
Index tests	<p>Index test: <u>TVUS</u> (TVS)</p> <p>Description of positive case definition by index test as reported: immobility of the rectum against the uterus and the posterior vaginal fornix considered as 'sliding sign negative', reflecting possible adhesion and endometriotic involvement of these structures; referenced to diagnostic criteria for all inspected sites of DIE: POD, USL, urinary bladder, RS and vagina</p> <p>Examiners: all TVS scans performed by 1 experienced examiner who was not blinded to clinical data</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: RS DIE</p> <p>Prevalence of target condition in the sample: pelvic peritoneum endometriosis 62/117 (53%), RS DIE 34/117 (29%)</p> <p>Reference standard: laparoscopy 117/117 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: DIE defined as subperitoneal endometriotic infiltration of tissues > 5 mm; histological presence of endometriosis taken to represent a 'true positive' diagnosis of endometriosis - histological criteria not specified; surgical procedure described in detail</p> <p>Examiners: 2 experienced surgeons not blinded to TVS results, surgical and pathological diagnostic criteria described</p>
Flow and timing	<p>Time interval between index test and reference standard: within 2 months</p> <p>Withdrawals: 25 patients excluded because they did not meet the inclusion criteria: 16 had a history of previous surgery for DIE, 3 had a history of gynaecological cancer, 4 women were virgins and 2 women did not provide consent</p>
Comparative	

Hudelist 2013 (Continued)

Key conclusions by the authors	TVS is a more useful test than vaginal examination for detecting endometriosis in the ovaries and rectosigmoid, and can be recommended as the method of choice for primary and preoperative assessment of patients with pelvic pain with suspected endometriosis
Conflict of interests	Not reported; supported by the OEGEO, Österreichische Gesellschaft für Endokrinologische Onkologie
Notes	Reported accuracy estimates for TVS for diagnosis of RS DIE confirmed as accurate

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Hudelist 2013 (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests?

No

High

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

Yes

Were all patients included in the analysis?

Yes

Low

Leon 2014
Study characteristics

Patient sampling

Primary objectives: to assess the performance of extended transvaginal sonography in diagnosing DIE, with surgical diagnostic laparoscopy and histological analysis used as the reference standard to confirm endometriosis

Study population: women with clinical suspicion of DIE based on clinical symptoms (chronic pelvic pain, deep dyspareunia, dyschezia, catamenial rectal bleeding, catamenial hematuria) or physical pelvic examination findings (non-mobile uterus, posterior vaginal fornix nodules, a painful pelvic examination)

Selection criteria: Inclusion criteria: clinical suspicion of DIE, patient's acceptance to undergo transvaginal sonography. Exclusion criteria: concomitant cancer, pregnancy, or pelvic inflammatory process; surgery performed at a centre other than the recruitment centre; choice of medical treatment instead of surgery; patient withdrawal before surgery

Study design: prospective, observational; non-consecutive enrolment

Patient characteristics and setting

Clinical presentation: dysmenorrhoea 51/51, dyspareunia 39/51, dyschezia 34/51, chronic pelvic pain 46/51, hematochezia 5/51; suspicious bimanual vaginal examination 26/51

Age: mean 32.9 ± 4.7 years, range 23 to 43 years

Number enrolled: 110 women

Number available for analysis: 51 women

Setting: Department of Obstetrics and Gynecology, Ultrasound and Human Reproduction Unit of the Indisa Clinic

Place of study: Santiago, Chile

Period of study: August 2011 to October 2012

Language: English

Leon 2014 (Continued)

Index tests	<p>Index test: <u>TVUS</u> (extended method: combination of bowel preparation with transvaginal gel instillation and use of 'sliding sign' for diagnosis)</p> <p>Description of positive case definition by index test as reported: DIE suspected in the presence of hypoechoic nodules at any area assessed. POD obliteration considered when the sliding sign was considered negative (described in details). Procedure and patients' preparation clearly described</p> <p>Examiners: all extended transvaginal sonographic examinations performed by 1 operator who had more than 10 years' experience in gynaecological sonography and 3 years' experience in assessment of deep infiltrating endometriosis; unclear whether operator was blinded to clinical data</p> <p>Interobserver variability: not provided</p>		
Target condition and reference standard(s)	<p>Target condition: DIE - separate anatomical sites</p> <p>Prevalence of target condition in the sample: DIE 39/51 (77%), POD obliteration 27/39 (69%)</p> <p>Reference standard: laparoscopy 51/51 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: surgeon asked to assess all areas under evaluation by sonography and to specifically determine the presence or absence of endometriotic implants. Any suspicious lesion should be biopsied for histological confirmation of the diagnosis. Procedure and diagnostic criteria not described</p> <p>Examiners: all patients underwent laparoscopy by 1 surgeon who was an expert in endometriotic surgery and was aware of index test results</p>		
Flow and timing	<p>Time interval between index test and reference standard: within 2 months (communication with study author)</p> <p>Withdrawals: 59 (54%) patients excluded because surgery was not performed at the centre of recruitment (n = 30), medical treatment was given instead of surgery (n = 25) and patients withdrew before surgery (n = 4)</p>		
Comparative			
Key conclusions by the authors	In conclusion, our results confirm that extended transvaginal sonography is useful for diagnosis of deep infiltrating endometriosis in advanced cases. Systematised use of sonography in compartments, bowel preparation and use of intravaginal gel together achieve optimal assessment of these patients, and we believe that this procedure is a useful and easy way to conduct a preoperative study in daily practice. Larger trials are needed to confirm our results		
Conflict of interests	Not reported		
Notes	<p>Reported accuracy estimates for extended TVUS with bowel preparation and intravaginal gel instillation for diagnosis of POD obliteration confirmed</p> <p>Accuracy estimates for bladder endometriosis reported by study authors but not presented in the review because this was not an assessed target condition</p> <p>Accuracy estimates for other anatomical sites of DIE (RVS, vaginal wall, RS) reported by study authors but not presented in the review because accuracy estimates were calculated per number of lesions, not per number of patients</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Leon 2014 (Continued)

Was a consecutive or random sample of patients enrolled?

No

Did the study avoid inappropriate exclusions?

Yes

Was a 'two-gate' design avoided?

Yes

High

Low

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard?

Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?

Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session?

Yes

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?

Unclear

Unclear

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?

Yes

Were the reference standard results interpreted without knowledge of the results of the index tests?

No

High

Low

DOMAIN 4: Flow and Timing

Leon 2014 (Continued)

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
High	

Manganaro 2012a
Study characteristics

Patient sampling	<p>Primary objectives: to define the role of 3 T Magnetom system MRI in evaluation of endometriosis</p> <p>Study population: women with clinical ± sonographic suspicion of endometriosis</p> <p>Selection criteria: Inclusion criteria: transvaginal ultrasound examination positive for endometriosis; patients with chronic pelvic pain; symptomatic patients with negative ultrasound; infertile patients</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: chronic pelvic pain, infertility; transvaginal ultrasound suggestive of endometriosis 23/46; treatment with combined oral contraceptive pill 17/46</p> <p>Age: mean 30.4 years, range 20 to 43 years</p> <p>Number enrolled: 46 women</p> <p>Number available for analysis: 46 women</p> <p>Setting: University Hospital: Umberto I Hospital, Sapienza University of Rome</p> <p>Place of study: Rome, Italy</p> <p>Period of study: February 2010 to September 2010</p> <p>Language: English</p>
Index tests	<p>Index test: <u>MRI</u> (3.0T Magnetom system (3.0T MRI))</p> <p>Description of positive case definition by index test as reported: imaging analysis performed using LMDSony 2451-MD monitor (resolution of 1220 × 1920 pixels); the following parameters were assessed: macroscopic endometriosis implants, deep endometriosis implants, fallopian tube involvement, presence of adhesions, fluid effusion in Douglas pouch, uterus and kidney pathologies, sacral nervous routes; endometriomas characterised as hyperintense in T1-w sequences, variable intensity in T2-w sequences, known as 'shading'; deep endometriosis implants characterised by both fibrosis components with low signal intensity in T2- and T1-w sequences, or for hyperintense foci in T1-w sequences and hypointense signal in T2-w; adhesions characterised as stellate or nodular hypointense areas on FSE T2-w sequences, creating attraction of close structures; rectosigmoid involvement showed increased thickness; rectouterine ligaments were thicker and were not homogeneous; unclear whether prespecified criteria or description of findings</p>

Manganaro 2012a (Continued)

Examiners: 2 radiologists with, respectively, 10 years' and 5 years' experience in female pelvis imaging; blinding to clinical data not reported

Interobserver variability: not provided; all images evaluated by 2 radiologists in consensus

Target condition and reference standard(s)

Target condition: pelvic endometriosis, DIE, ovarian endometriosis

Prevalence of target condition in the sample: pelvic endometriosis 32/46 (69.6%), DIE 23/46 (50%), ovarian endometriosis 19/46 (41%)

Reference standard: laparoscopy 46/46 (100%)

Description of positive case definition by reference test as reported: not specified; procedure reported as 'Diagnostic and therapeutic laparoscopy', not described

Examiners: numbers or level of expertise of surgeons not provided; unclear whether blinded to results of the index test

Flow and timing

Time interval between index test and reference standard: < 12 months (communication with study authors)

Withdrawals: none reported

Comparative

Key conclusions by the authors

Pelvic MRI performed with 3T system guarantees high spatial and contrast resolution, providing accurate information about endometriosis implants, with good presurgery mapping of lesions involving both bowels and bladder surface and rectouterine ligaments

Conflict of interests

Not reported

Notes

Reported accuracy estimates for 3.0T MRI for diagnosis of pelvic, ovarian and DIE endometriosis were calculated on the basis of data presented by study authors in personal communication

Possible partial overlap with another study from the same group (Manganaro 2013 (study period July 2010 to July 2012)); however, different sites of endometriosis assessed, and both studies included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge	Yes		

Manganaro 2012a *(Continued)*
of the results of the reference
standard?

Did the study provide a clear pre-specified definition of what was considered to be a “positive” result of index test?	Unclear
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear
Unclear Low	

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Unclear
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Unclear Low	

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Low	

Manganaro 2012b
Study characteristics

Patient sampling	Primary objectives: to define the role of 3T Magnetom system MRI in the evaluation of posterior cul-de-sac obliteration in endometriosis Study population: women with clinical ± sonographic suspicion of endometriosis
------------------	--

Manganaro 2012b (Continued)

	<p>Selection criteria: Inclusion criteria: transvaginal ultrasound examination positive for endometriosis; patients with chronic pelvic pain; symptomatic patients with negative ultrasound; infertile patients</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: transvaginal ultrasound examination positive for endometriosis, chronic pelvic pain, symptomatic patients with negative ultrasound examination</p> <p>Age: mean 26 years, range 19 to 35 years</p> <p>Number enrolled: 19 women</p> <p>Number available for analysis: 19 women</p> <p>Setting: University Hospital: Umberto I Hospital, Sapienza University of Rome</p> <p>Place of study: Rome, Italy</p> <p>Period of study: October 2010 to April 2011</p> <p>Language: English</p>
Index tests	<p>Index test: MRI (3.0T Magnetom system (3.0T MRI))</p> <p>Description of positive case definition by index test as reported: imaging analysis using LMDSony 2451-MD monitor (resolution of 1220 × 1920 pixels); assessment of 3 pelvic compartments - anterior, medium, posterior - for the following: macroscopic endometriosis implants (> 5 mm), adhesions (disappearance of the fat tissue plane separating different structures), USL involvement (increased and inhomogeneous thickness and abnormal arciform appearance), PCS obliteration (retroflexed uterus, tethered appearance of the rectum in the direction of the uterus, strands between uterus and intestine, fibrotic plaque covering the serosal surface of the uterus and elevated posterior cervical fornix) and signal intensity of endometriotic lesions</p> <p>Examiners: 2 radiologists with 12 years' and 7 years' experience in female pelvis imaging; blinded to clinical data</p> <p>Interobserver variability: validated analysis showed k value of 0.72 with substantial degree between 2 radiologists</p>
Target condition and reference standard(s)	<p>Target condition: POD obliteration by endometriosis</p> <p>Prevalence of target condition in the sample: POD obliteration 15/19 (79%)</p> <p>Reference standard: laparoscopy 19/19 (100%)</p> <p>Description of positive case definition by reference test as reported: not specified; procedure reported as "Diagnostic and therapeutic laparoscopy", endometriotic lesions mapped and staging established</p> <p>Examiners: numbers or level of expertise of surgeons not provided; unclear whether blinded to results of the index test</p>
Flow and timing	<p>Time interval between index test and reference standard: within 3 months</p> <p>Withdrawals: none reported</p>
Comparative	
Key conclusions by the authors	<p>Precise preoperative mapping of posterior cul-de-sac region is essential for preoperative planning. In our work, 3-T MRI was shown to be excellent in evaluation of posterior cul-de-sac obliteration associated with optimal evaluation of uterosacral ligaments due to higher-contrast spatial resolution</p>

Manganaro 2012b (Continued)

Conflict of interests	All study authors have no conflict of interest or financial relationship to disclose
Notes	Reported accuracy estimates for 3.0T MRI for diagnosis of POD obliteration confirmed as accurate Estimates for diagnosis of USL not reported as presented in larger overlapping study (Manganaro 2013)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without	Unclear		

Manganaro 2012b (Continued)
knowledge of the results of the
index tests?

	Unclear	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Manganaro 2013

Study characteristics

Patient sampling	<p>Primary objectives: to assess diagnostic accuracy of 3.0T magnetic resonance imaging (3.0T MRI) in assessing involvement of uterosacral ligaments (USLs) in deep infiltrating endometriosis (DIE)</p> <p>Study population: patients with suspected USL DIE based on clinical symptoms, abnormal gynaecological examination or transvaginal ultrasound findings</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: severe pain symptoms such as dyspareunia, dysmenorrhoea and acyclical pain (visual analogue scale (VAS) > 7/10)</p> <p>Age: mean 28 years, range 19 to 45 years</p> <p>Number enrolled: 42 women</p> <p>Number available for analysis: 42 women</p> <p>Setting: University Hospital, Umberto I Hospital, "Sapienza" University of Rome</p> <p>Place of study: Rome, Italy</p> <p>Period of study: July 2010 to July 2012</p> <p>Language: English</p>
Index tests	<p>Index test: MRI (3.0T MRI)</p> <p>Description of positive case definition by index test as reported: USL DIE diagnosed in the presence of 1 of the following: asymmetry between the 2 ligaments; increased and non-homogeneous thickness associated with abnormal arciform appearance (> 3 mm) and/or a nodule (> 5 mm) with regular or irregular stellate margins; thickening of the posterior uterine wall associated with lower signal intensity; loss of the fatty plane between uterus and rectum on T2-w longitudinal images, indicating adhesions; change in signal intensity in USL area (hypointense signal in both T1-/T2-w sequences with hyperintense foci on T2 sequences, which may indicate fibrosis with glandular spots,</p>

Manganaro 2013 (Continued)

or hypointense signal on T1-/T2-weighted images with hyperintense foci on T1-weighted owing to hemorrhagic foci within fibrotic issue and/or hypointense signal in both T1-/T2-w sequences if fibrotic reaction is abundant)

Examiners: radiologist who analysed images had > 13 years' experience in imaging of the female pelvis (single operator) and was blinded to results of previous imaging or clinical examination

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: USL DIE

Prevalence of target condition in the sample: pelvic endometriosis 42/42 (100%), USL DIE 19/42 (45.2%)

Reference standard: laparoscopy 42/42 (100%) + histopathology

Description of positive case definition by reference test as reported: visual inspection (rASRM classification) and histopathological examination - histological criteria not specified; procedure described

Examiners: laparoscopy performed by the same surgeon with more than 20 years' experience; unclear whether blinded to results of the index test; number and level of expertise of pathologists not reported

Flow and timing

Time interval between index test and reference standard: within 3 months

Withdrawals: none reported

Comparative

Key conclusions by the authors

In conclusion, we can assume that 3.0T MRI is an optimal technique that may enable accurate preoperative assessment to select patients for the correct kind of surgery or follow-up

Conflict of interests

Study authors declared no conflict of interest

Notes

Reported accuracy estimates for 3.0T MRI for diagnosis of USL endometriosis confirmed as accurate 100% of study participants diagnosed with ovarian endometrioma, likely that this is a highly preselected group; however, radiologist performing index test was blinded to clinical data

Possible partial overlap with another study from the same group (Manganaro 2011 (study period February 2010 to September 2010)); however, different sites of endometriosis assessed, and both studies included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	Low

Manganaro 2013 (Continued)

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a “positive” result of index test? Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice? Yes

Low
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Unclear
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Mangler 2013

Study characteristics

Patient sampling	<p>Primary objectives: to identify the sensitivity and specificity of different diagnostic procedures in preoperatively assessing bowel infiltration among patients with rectovaginal endometriosis, including rectovaginal gynaecological examination, transvaginal sonography, MRI or rectal endosonography combined with rectosigmoidoscopy</p> <p>Study population: patients with suspected/known rectovaginal endometriosis who were operated on at the study authors' institution. Endometriosis suspected on the basis of clinical symptoms, abnormal gynaecological examination or other imaging tests, or known through previous operations</p> <p>Selection criteria: not specified</p> <p>Study design: observational; prospective collection of data; consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 73%, bowel symptoms (dyschezia, cyclical constipation, diarrhoea) 68%; overall 97% presented with symptoms; previous surgery for pelvic pain 78%; hormonal treatment 69%</p> <p>Age: mean 34 years, range 19 to 51 years</p> <p>Number enrolled: 79 women</p> <p>Number available for analysis: 79 women</p> <p>Setting: University Hospital, Charité Campus Mitte</p> <p>Place of study: Berlin, Germany</p> <p>Period of study: September 2007 to February 2010</p> <p>Language: English</p>
Index tests	<p>Index test: TVUS(vaginal ultrasound)</p> <p>Description of positive case definition by index test as reported: 'During the transvaginal ultrasound, special focus was given to the rectovaginal septum and adjacent bowel'; diagnostic criteria and procedure not described</p> <p>Examiners: consultants who were not aware of results of the other tests and of the reference procedure</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: RS DIE ('bowel-infiltrating rectovaginal endometriosis')</p> <p>Prevalence of target condition in the sample: RS DIE endometriosis 48/79 (61%), overall endometriosis 79/79 (100%)</p> <p>Reference standard: surgery (vaginal approach + laparoscopy ± laparotomy) 79/79 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: procedure and diagnostic criteria described (referenced to primary source (Mangler et al., 2008))</p> <p>Examiners: all operations carried out by a single consultant, who had access to preoperative findings</p>
Flow and timing	<p>Time interval between index test and reference standard: within 2 to 6 weeks</p> <p>Withdrawals: none reported in ultrasound group</p>
Comparative	

Mangler 2013 (Continued)

Key conclusions by the authors	In conclusion, as a result of the dilemma of preoperative imaging, we propose that the standardised vaginal operative approach be used to verify the extent of endometriosis in the rectovaginal septum until valid imaging techniques become available to accurately assess preoperatively the growth of rectovaginal endometriosis
Conflict of interests	Study authors declared no conflict of interest
Notes	<p>Reported accuracy estimates for TVUS for diagnosis of RS endometriosis confirmed as accurate</p> <p>100% of study participants diagnosed with endometriosis, likely that this is a highly preselected group; however, radiologist performing index test blinded to clinical data</p> <p>Data on diagnostic performance of pelvic examination and Ca-125 also presented - not included as beyond the scope of this review</p> <p>Data on diagnostic performance of rectal endosonography, rectosigmoidoscopy and MRI also presented - not included, as information is insufficient for construction of 2 × 2 tables and raw data were not available from study authors</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	No		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	No		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		

Mangler 2013 (Continued)

	High	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	No	
	High	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Menada 2008a

Study characteristics	
Patient sampling	<p>Primary objectives: to compare the effectiveness of transvaginal ultrasonography (TVS) and water contrast in the rectum during transvaginal ultrasonography (RWC-TVS) for the diagnosis of rectal infiltration in women with rectovaginal endometriosis</p> <p>Study population: women with suspected rectovaginal endometriosis on the basis of pain symptoms and/or gynaecological examination</p> <p>Selection criteria: exclusion criteria: patients who were virgins or who had any type of genital malformation that made physical examination or TVS impossible; previous surgical excision of bowel endometriosis</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 84/90, dyspareunia 68/90, chronic pelvic pain 62/90, infertility 32/90, diarrhoea and/or constipation 61/90, bowel movement pain or cramping 69/90, pain on defecation 32/90, rectal bleeding 16/90, lower back pain 57/90; previous medical treatments for endometriosis 82/90</p> <p>Age: median 32 years, range 18 to 42 years</p> <p>Number enrolled: 90 women</p> <p>Number available for analysis: 90 women</p> <p>Setting: University Hospital, San Martino Hospital, University of Genoa</p>

Menada 2008a (Continued)

Place of study: Genoa, Italy

Period of study: October 2006 to November 2007

Language: English

Index tests	Index test: TVUS 2 types (TVS; RWC-TV) Description of positive case definition by index test as reported: rectovaginal endometriosis appearing ultrasonographically as rounded or triangular hypoechoic masses, located anterior or lateral to the rectum, immediately adjacent or close to the rectal wall. Rectal endometriotic infiltration defined by the fact that the rectovaginal hypoechoic mass was adherent and/or penetrated into the intestinal wall thickening of the muscularis mucosa; hypoechoic or hyperechoic foci sometimes present Examiners: 2 different experienced ultrasonographers independently performed examinations: 1 operator performed all TVS, second operator performed RWC-TV. Operators were informed that rectovaginal endometriosis was suspected, but they were not aware of the findings of vaginal or rectal examination, and they were not informed of the findings of previous radiological examinations and results of other index tests Interobserver variability: not provided		
Target condition and reference standard(s)	Target condition: RVS endometriosis (rectovaginal endometriosis) Prevalence of target condition in the sample: pelvic endometriosis 81/90 (90%), rectovaginal endometriosis 69/90 (76.7%), rectal infiltration 29/90 (32.2%) Reference standard: laparoscopy, laparotomy (number in each group not specified) 90/90 (100%) + histopathology Description of positive case definition by reference test as reported: visual diagnosis confirmed by histopathology; criteria (referenced to a primary source) described; surgical procedure described in detail Examiners: all surgical procedures performed by a team of gynaecological and colorectal surgeons with extensive experience in the treatment of pelvic and bowel endometriosis; unclear whether blinded to results of the index tests; numbers and level of expertise of pathologists not reported		
Flow and timing	Time interval between index test and reference standard: within several hours Withdrawals: none reported		
Comparative			
Key conclusions by the authors	RWC-TV determines the presence of rectovaginal nodules infiltrating the rectal muscularis propria more accurately than TVS; RWC-TV could be used when TVS cannot exclude the presence of rectal infiltration		
Conflict of interests	Not reported		
Notes	Reported accuracy estimates for TVS and RWC-TV for diagnosis of rectovaginal endometriosis confirmed as accurate Data on the tolerability of each of the index tests and on comparison between index tests for diagnosis of different types of lesions also presented - not included in this review		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Menada 2008a *(Continued)*

Was a consecutive or random sample of patients enrolled?	No	
Did the study avoid inappropriate exclusions?	Yes	
Was a 'two-gate' design avoided?	Yes	
	High	Low

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes	
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes	
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes	
	Low	Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes	
--	-----	--

Menada 2008a (Continued)

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Okada 1995
Study characteristics

Patient sampling	<p>Primary objectives: to assess the usefulness of fat-saturated magnetic resonance imaging (MRI) in detecting small endometrial implants by comparing it with conventional MRI</p> <p>Study population: women visiting outpatient department with suspected endometriosis based on clinical presentation (symptoms and pelvic examination), transvaginal ultrasonography and/or blood test for Ca-125</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: infertility, lower abdominal pain, menstrual pain, dyspareunia; suspected endometriosis on pelvic examination or transvaginal ultrasonography</p> <p>Age: mean 37.4 years, range 26 to 49 years</p> <p>Number enrolled: 74 women</p> <p>Number available for analysis: 74 women</p> <p>Setting: University Hospital, Shimane Medical University</p> <p>Place of study: Izumo, Japan</p> <p>Period of study: August 1991 to December 1993</p> <p>Language: Japanese</p>
Index tests	<p>Index test: <u>MRI</u> (T1-w fat-saturated MRI)</p> <p>Description of positive case definition by index test as reported: diagnosis of endometriosis through MRI based on previously published criteria (referenced to Togashi et al., 1991) described</p> <p>Examiners: numbers of operators, level of expertise or blinding to clinical data not reported</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: pelvic endometriosis</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 65/74 (87.8%); stage I 11/65 (17%); stage II 7/65 (11%); stage III 11/65 (17%); stage IV 36/65 (55%)</p> <p>Reference standard: laparoscopy 47/74 (63.5%), laparotomy 27/74 (36.5%) + histopathology</p> <p>Description of positive case definition by reference test as reported: visual inspection of pelvic cavity (AFS classification and General Rules for Clinical Management of Endometriosis, Japan Society of Obstetrics and Gynecology); histopathological examination when available, histological criteria included: 1 - endometrial epithelial cells and endometrial stromal cells; 2</p>

Okada 1995 (Continued)

- haemorrhages in endometrial glandular cavities or in surrounding interstitium; 3 - pigmentation; 4 - macrophage phagocytosis; surgical procedure described

Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinded to results of the index test

Flow and timing	Time interval between index test and reference standard: within 2 weeks Withdrawals: none reported
Comparative	
Key conclusions by the authors	Fat-saturated MRI can be used for detecting small endometrial implants
Conflict of interests	Not reported
Notes	Reported summary statistics of fat-saturated MRI for diagnosis of pelvic endometriosis confirmed as accurate Accuracy estimates for different types/sizes of lesions also presented – not included in this review Accuracy estimates for conventional MRI calculated by number of lesions, not by number of patients - not presented in this review Likely overlap with 2 smaller studies from the same group (Takahashi 1994 (study period not specified) and Sigumura 1993 (study period May 1991 to August 1992)) - Not able to contact study authors; therefore, data from Sigumura 1993 presented for conventional MRI only; Takahashi 1994 excluded

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		

Okada 1995 (Continued)

Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear	
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear	
		Unclear
		Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
		Unclear
		Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
		Low

Pascual 2010

Study characteristics		
Patient sampling	<p>Primary objectives: to evaluate the diagnostic accuracy of Introital 3-dimensional ultrasound (Introital 3D-US) in the identification of rectovaginal septum endometriosis</p> <p>Study population: patients with clinically suspected endometriosis based on patient history of pelvic pain and/or clinical examination</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational; consecutive enrolment</p>	
Patient characteristics and setting	<p>Clinical presentation: dyspareunia and/or dysmenorrhoea 39/39, infertility 15/39; previous treatment for persistent pelvic pain with estrogens, progestins and/or GnRH agonist and non-steroidal anti-inflammatory drugs for ≥ 1 year 39/39</p>	

Pascual 2010 (Continued)

Age: mean 35.6 ± 5.7 years, range 25 to 44 years

Number enrolled: 39 women

Number available for analysis: 38 women

Setting: University Hospital, Instituto Universitario Dexeus of Barcelona

Place of study: Barcelona, Spain

Period of study: January 2008 to July 2009

Language: English

Index tests

Index test: TVUS (Introital 3D-US)

Description of positive case definition by index test as reported: data file sent via Digital Imaging and Communication in Medicine to a personal computer and stored to be analysed with use of appropriate software; deep endometriosis implants suspected by the presence of hypoechoic areas, nodules or anatomical distortion of this specific location with use of render mode in the coronal plane obtained after multi-planar navigation; unclear whether prespecified criteria or description of findings

Examiners: transvaginal US scans carried out by 3 experienced examiners, using the same scanning protocol; stored 3D volumes analysed by just 1 examiner; unclear whether blinded to clinical data

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: RVS endometriosis (deep rectovaginal septum endometriosis)

Prevalence of target condition in the sample: pelvic endometriosis 38/38 (100%), deep rectovaginal septum endometriosis 19/38 (50%)

Reference standard: laparoscopy 38/38 (100%) + histopathology

Description of positive case definition by reference test as reported: visual diagnosis (AFS classification); diagnosis of rectovaginal endometriosis proved histologically for each patient - criteria not specified; surgical procedure explained in detail

Examiners: numbers or level of expertise of surgeons or pathologists not provided; unclear whether blinded to results of the index test

Flow and timing

Time interval between index test and reference standard: within 1 month

Withdrawals: in 1 case (3%), volume quality not adequate to be re-elaborated because of poor visualisation of the rectum - this case was not considered in the statistical analysis

Comparative

Key conclusions by the authors

Introital 3D ultrasonography seems an effective method for diagnosis of endometriosis of the rectovaginal septum and should be included in the preoperative evaluation of patients with clinical suspicion of deep endometriosis

Conflict of interests

Not reported

Notes

Reported accuracy estimates for introital 3D US for diagnosis of RVS endometriosis confirmed as accurate

2D-US reported as part of the study protocol, but no data available for 2 × 2 tables - not included in this review

Pascual 2010 (Continued)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Unclear		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Unclear		
		Unclear	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Pascual 2010 (Continued)

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Piessens 2014
Study characteristics

Patient sampling **Primary objectives:** to assess whether diagnostic accuracy for DIE comparable with that of the experts could be obtained and used to assess the learning curve associated with acquiring these skills

Study population: patients with clinically suspected endometriosis referred to TVUS

Selection criteria: not specified

Study design: observational; prospective consecutive enrolment; retrospective analysis

Patient characteristics and setting **Clinical presentation:** dysmenorrhoea (63%), dyschezia (53%), dyspareunia (44%), infertility (22%), abnormal bleeding (20%), chronic pain (21%), rectal bleeding (8%); past history of endometriosis (72%)

Age: range 18 to 48 years

Number enrolled: 205 women

Number available for analysis: 85 women

Setting: Monash Health, Clayton; Monash University

Place of study: Clayton Victoria, Australia

Period of study: November 2009 to September 2011

Language: English

Index tests **Index test:** TVUS-BP (DIE-TVUS)

Description of positive case definition by index test as reported: DIE suspected in the presence of hypoechoic nodules at any area assessed; POD obliteration considered when the sliding sign was considered negative (described and referenced). Examination protocol described; all examinations interpreted in real time and recorded on DVD

Examiners: all examinations performed by a single operator who is a gynaecologist with a subspecialty degree in ultrasound and more than 10 years' experience, but no prior experience in detecting DIE; operator was not blinded to symptoms and history of women

Interobserver variability: not provided

Target condition and reference standard(s) **Target condition:** DIE at specific anatomical sites, ovarian endometrioma

Prevalence of target condition in the sample: bowel endometriosis 24/85 (7%), POD obliteration 34 (40%), vaginal endometriosis 15/85 (18%), ovarian endometrioma 17/85 (20%)

Reference standard: laparoscopy 85/85 (100%) + histopathology

Piessens 2014 (Continued)

Description of positive case definition by reference test as reported: visual diagnosis at laparoscopy; lesions not confirmed by histology were not excluded - criteria not specified; surgical procedure not explained

Examiners: numbers or level of expertise of surgeons or pathologists not provided; unclear whether blinded to results of the index test. Laparoscopic images/reports assessed by study author, who was blinded to the ultrasound report at the time of analysis

Flow and timing	<p>Time interval between index test and reference standard: within 12 months (personal communication with study author)</p> <p>Withdrawals: 120 (59%) patients did not undergo surgery (reason not specified) and were excluded from the analyses</p>
Comparative	
Key conclusions by the authors	In conclusion, this study provides evidence that the skills required to diagnose DIE can be attained by experienced sonologists after a brief learning period. The study provides further external validation of the high diagnostic accuracy of DIE-TVUS
Conflict of interests	Not reported
Notes	<p>Reported diagnostic accuracy estimates for TVUS for specific sites of DIE and endometrioma confirmed as accurate</p> <p>Accuracy estimates for bladder endometriosis reported by study authors but not presented in the review because they were not the assessed target conditions</p> <p>Learning curve assessed using a validated statistical model and reported - not included in the review</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		

Piessens 2014 *(Continued)*

Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	No
High	
Low	
DOMAIN 3: Reference Standard	
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
High	
Low	
DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
High	

Piketty 2009

Study characteristics	
Patient sampling	<p>Primary objectives: to compare results of transvaginal ultrasonography (TVUS) with those of transrectal ultrasonography (TRUS), which is the investigation of choice for the diagnosis of rectal wall infiltration, and to clarify whether TVUS is limited</p> <p>Study population: patients suffering from pelvic pain (alone or associated with infertility) who underwent complete surgical exeresis of deeply infiltrating endometriosis (DIE), which was suspected in all cases preoperatively (questioning, clinical examination, imaging)</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>

Piketty 2009 (Continued)

Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea, deep dyspareunia, non-cyclical chronic pelvic pain, gastrointestinal symptoms, lower urinary tract symptoms; previous hormonal treatment for endometriosis 134/134, previous surgery for endometriosis 88/134</p> <p>Age: mean 32.1 ± 5.0 years, range 22 to 47 years</p> <p>Number enrolled: 134 women</p> <p>Number available for analysis: 134 women</p> <p>Setting: University Hospital, Université Paris Descartes</p> <p>Place of study: Paris, France</p> <p>Period of study: January 2005 to July 2007</p> <p>Language: English</p>
Index tests	<p>Index test: <u>TVUS</u>; <u>TRUS</u></p> <p>Description of positive case definition by index test as reported: <u>TVUS</u> - DIE defined as presence of hypoechoic and irregular nodes in assessed pelvic structures; intestinal DIE (ileum - rectum) defined as previously published (referenced to Bazot et al., 2007) and described; <u>TRUS</u> - DIE showed up as hypoechoic peridigestive nodules of rounded or roughly triangular shape (ileum - rectum); diagnosis of bowel infiltration in accordance with previously published (referenced to Chapron et al., 1998) and described</p> <p>Examiners: <u>TVUS</u> - single experienced radiologist; <u>TRUS</u> - single examiner, level of expertise not reported; both examiners informed that DIE was suspected but blinded to the results of clinical findings and previous imaging examinations</p> <p>Interobserver variability: not provided</p>
Target condition and reference standard(s)	<p>Target condition: bowel (ileum - rectum) endometriosis</p> <p>Prevalence of target condition in the sample: DIE 134/134 (100%), bowel endometriosis 75/134 (56%)</p> <p>Reference standard: laparoscopy, laparotomy (numbers for each procedure not specified) + histopathology</p> <p>Description of positive case definition by reference test as reported: diagnosis based on histological assessment - criteria not specified; surgical procedure not described</p> <p>Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinding to the results of index tests</p>
Flow and timing	<p>Time interval between index test and reference standard: < 12 months (communication with study authors)</p> <p>Withdrawals: no withdrawals reported for TRUS, 1 (1%) unexplained withdrawal in TVUS group</p>
Comparative	
Key conclusions by the authors	<p>TVUS and TRUS have similar degrees of accuracy for predicting intestinal involvement. TVUS must be the first-line imaging process performed for patients presenting with clinically suspected DIE. The question for the coming years is to define whether it is necessary for TRUS to be carried out systematically in cases of clinically suspected DIE</p>
Conflict of interests	Not reported
Notes	Reported accuracy estimates for TVUS and TRUS for diagnosis of bowel DIE confirmed as accurate

Piketty 2009 (Continued)

Other sites of DIE assessed, but data are sufficient for calculation of accuracy estimates for these areas - not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low

Piketty 2009 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Low	

Reid 2013a
Study characteristics

Patient sampling	<p>Primary objectives: to evaluate real-time dynamic transvaginal sonography (TVS) for the prediction of pouch of Douglas (POD) obliteration in women undergoing laparoscopy for suspected endometriosis; specifically, real-time dynamic TVS used to evaluate the sliding sign technique to predict POD obliteration</p> <p>Study population: women with a history of chronic pelvic pain and/or endometriosis and scheduled for operative laparoscopy</p> <p>Selection criteria: Inclusion criteria: pelvic pain, defined as chronic if it persisted for longer than 3 months and could be constant or intermittent, cyclical or non-cyclical in nature; 4 types of pelvic pain included: cyclical pain during menstruation (dysmenorrhoea), deep dyspareunia, dyschezia and non-cyclical pelvic pain; only women of reproductive age. Exclusion criteria: malignancy, menopause, pregnancy</p> <p>Study design: multi-centre, prospective, observational study; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: cyclical pain 70/100, pain requiring strong analgesia 49/100, pain affecting life despite strong analgesia 53/100, pain preventing daily activities 55/100, dyspareunia 56/100, dyschezia 51/100, tenesmus 29/100, cyclical constipation 32/100, cyclical diarrhoea 37/100 (37%), cyclical hematuria 3/100 (3%), cyclical hematochezia 16/100 (16%), constant pain 2/100 (2%), non-cyclical pain 2/100; pain location: left iliac fossa pain 49%, lower abdominal pain 65%, right iliac fossa pain 44%, left upper quadrant pain 7%, epigastric pain 2%, right upper quadrant pain 2% and back pain 2%; median duration of pelvic pain 18 months; history of in vitro fertilisation (13%), irregular menstrual periods (19%), use of contraception (30%), history of infertility (30%) and history of endometriosis (60%)</p> <p>Age: mean 32.78 ± 6.28 years; median 33.0 years, range 19 to 48 years</p> <p>Number enrolled: 100 women? (see note below)</p> <p>Number available for analysis: 100 women</p> <p>Setting: 4 university teaching hospitals, tertiary referral centres: Nepean Hospital, Royal Hospital for Women, Royal Prince Alfred Hospital, Liverpool Hospital; 5 private hospitals: Norwest Private Hospital, Hurstville Private Hospital, St. Luke's Private Hospital, Prince of Wales Private Hospital, St. George Private Hospital</p> <p>Place of study: NSW, Australia</p> <p>Period of study: January 2009 to November 2011</p> <p>Language: English</p>
Index tests	Index test: TVUS, sliding sign (TVS)

Reid 2013a (Continued)

Description of positive case definition by index test as reported: 'sliding sign' as marker of POD obliteration described: sliding sign considered positive if anterior rectal wall glided smoothly over posterior cervix and posterior vaginal wall, or if anterior rectosigmoid wall glided smoothly over posterior upper uterus/fundus; if positive 'sliding sign' in posterior cervix and posterior upper uterus, POD recorded as 'not obliterated'; if negative 'sliding sign' in either of these anatomical regions, POD recorded as 'obliterated'. Further TVS assessment included evaluation of uterus, ovaries and posterior compartment for DIE (rectosigmoid/anterior rectum, USL, RVS/vaginal) - criteria not stated

Examiners: single examiner; level of expertise and blinding to clinical data not reported

Interobserver variability: not provided

Target condition and reference standard(s)	Target condition: posterior DIE - separate anatomical sites Prevalence of target condition in the sample: pelvic endometriosis 84/100 (84%), posterior DIE 33/100 (33%) Reference standard: laparoscopy 100/100 (100%) + histopathology Description of positive case definition by reference test as reported: surgical diagnosis of endometriosis made in accordance with published criteria (referenced to Bazot et al., 2003) and described; surgical procedure not described Examiners: surgery performed by a total of 7 advanced laparoscopic surgeons, all of whom are experienced in excision of DIE; data on numbers or level of expertise of pathologists not provided; unclear whether blinded to results of the index test		
Flow and timing	Time interval between index test and reference standard: < 12 months (communication with study authors) Withdrawals: none reported		
Comparative			
Key conclusions by the authors	Preoperative real-time dynamic TVS evaluation using the sliding sign seems to establish with a high degree of certainty whether the POD is obliterated. Given the increased risk of deep infiltrating endometriosis among women with POD obliteration, the TVS sliding sign technique may be useful for identifying women who may be at higher risk for bowel endometriosis		
Conflict of interests	Not reported		
Notes	Reported accuracy estimates for TVS for diagnosis of specific sites of posterior DIE and for sliding sign for diagnosis of obliterated POD confirmed as correct Data on the overall posterior DIE as a group not available "Complete TVS sliding sign and laparoscopic data available for 100 women" raises concern about underreported number of enrolled participants; therefore, unclear whether any withdrawals not mentioned		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Reid 2013a *(Continued)*

Did the study avoid inappropriate exclusions?

Yes

Was a 'two-gate' design avoided?

Yes

High

Low

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard?

Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?

Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session?

Yes

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?

Yes

Low

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?

Yes

Were the reference standard results interpreted without knowledge of the results of the index tests?

Unclear

Unclear

Low

DOMAIN 4: Flow and Timing

Reid 2013a (Continued)

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

Yes

Were all patients included in the analysis?

Yes

Low

Reid 2014
Study characteristics

Patient sampling **Primary objectives:** to evaluate the use of a newly modified sonovaginography (SVG) technique, outpatient 'office gel SVG', for prediction of posterior compartment deep infiltrating endometriosis (DIE)

Study population: women who presented to pelvic pain clinic with symptoms suggestive of endometriosis

Selection criteria: Inclusion criteria: reproductive age, history of chronic pelvic pain \pm history of endometriosis, laparoscopy within 6 months of gel SVG examination. Exclusion criteria: malignancy, menopause, pregnancy

Study design: multi-centre, prospective, observational study; consecutive enrolment

Patient characteristics and setting **Clinical presentation:** chronic pelvic pain, dysmenorrhoea, dyspareunia, dyschezia; mean duration of pain 39.7 ± 47.5 months; history of infertility 44/220; history of endometriosis 92/220; history of bowel DIE in the past 10/220

Age: mean 32.2 ± 7.5 years

Number enrolled: 220 women

Number available for analysis: 189 women

Setting: 4 university teaching hospitals, tertiary referral centres: Nepean Hospital, Royal Hospital for Women, Royal Prince Alfred Hospital, Liverpool Hospital; 5 private hospitals: Norwest Private Hospital, Hurstville Private Hospital, St. Luke's Private Hospital, Prince of Wales Private Hospital, St. George Private Hospital

Place of study: NSW, Australia

Period of study: January 2009 to February 2013

Language: English

Index tests **Index test:** Sonovaginography (SVG)

Description of positive case definition by index test as reported: RVS DIE predicted during SVG when hyperechoic RVS layer was interrupted or was no longer visible. USL nodules identified as defined hypoechoic lesions located laterally alongside the cervix. Vaginal DIE identified as solid nodule in vaginal wall. Diagnosis of rectal/rectosigmoid DIE made when the normally appearing linear hypoechoic longitudinal muscle was thickened by a solid hypoechoic mass, which was continuous within the longitudinal muscle of the bowel and could be visualised in both sagittal and transverse planes. The 'sliding sign' was performed during gel SVG to determine whether adhesions existed between anterior rectum/rectosigmoid and posterior

Reid 2014 (Continued)

vaginal wall/cervix/uterus (i.e. POD obliteration). Description is supported by pictures of images for each compartment. Technique is described in detail

Examiners: all SVG examinations performed by 2 operators (1 was an expert gynaecological sonologist with experience in diagnosis of DIE; the other was a gynaecological ultrasound fellow supervised by an experienced operator). Same person who performed SVG performed the gynaecological examination and TVS. Operators were not blinded to clinical history

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: posterior DIE - overall and separate anatomical sites (USL, RVS, vagina, bowel including anterior rectum and rectosigmoid)

Prevalence of target condition in the sample: pelvic endometriosis 146/189 (77%), posterior DIE 57/189 (30%), separate compartments of endometriosis: bowel 43/189 (23%), vaginal 11/189 (6%), RVS 11/189 (6%), USL 10/189 (5%), POD obliteration 47/189 (25%)

Reference standard: laparoscopy 189/189 (100%) + histopathology

Description of positive case definition by reference test as reported: surgical diagnosis of endometriosis made if any of the following was satisfied: 1 - histological confirmation of endometriosis in ≥ 1 resected nodule; 2 - visualisation and palpation of subperitoneal nodule without biopsy and another histologically proven location of endometriosis; 3 - visualisation of complete obliteration of cul-de-sac. Surgical findings verified through primary author's review of detailed operation reports and diagrams made at the time of surgery; surgical procedure described

Examiners: surgery performed by a total of 13 laparoscopic surgeons: 9 advanced laparoscopic surgeons and 4 general gynaecological surgeons. Surgeons not blinded to patient data, including results of the index test

Flow and timing

Time interval between index test and reference standard: within 6 months

Withdrawals: 31 women chose not to undergo surgery after consultation with their surgeon; no additional details provided

Comparative

Key conclusions by the authors

Office gel SVG appears an effective imaging technique for detection of bowel DIE, with higher accuracy for prediction of rectosigmoid vs anterior rectum DIE. SVG found to have high specificity and NPV for all forms of DIE, indicating that negative SVG examination correlates highly with absence of DIE at laparoscopy. This new technique may not only aid in triaging of women for referral to an advanced laparoscopic surgeon \pm colorectal input, but may act as a useful learning tool for visualisation of posterior pelvic compartment in women with suspected DIE

Conflict of interests

Not reported

Notes

Reported accuracy estimates for SVG for diagnosis of posterior DIE, overall and specific sites (rectosigmoid/anterior rectum, USL, RVS/vagina, POD obliteration), confirmed as correct

Methodological quality

Item

Authors' judgement

Risk of bias

Applicability concerns

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?

Yes

Reid 2014 (Continued)

Did the study avoid inappropriate exclusions? Yes

Was a 'two-gate' design avoided? Yes

Low
Low
DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test? Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice? No

High
Low
DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? No

High
Low
DOMAIN 4: Flow and Timing

Reid 2014 (Continued)

Was there an appropriate interval between index test and reference standard?

Yes

Did all patients receive the same reference standard?

Yes

Were all patients included in the analysis?

Yes

Low

Ribeiro 2008a
Study characteristics

Patient sampling

Primary objectives: to evaluate the sensitivity, specificity, NPV, PPV, association and agreement of double-contrast barium enema (DCBE) and transrectal endoscopic ultrasonography (TrEUS) in the diagnosis of rectosigmoid colon endometriosis

Study population: patients with clinically suspected deeply infiltrating endometriosis (DIE) referred to gynaecological endoscopy and endometriosis clinic

Selection criteria: Inclusion criteria: dysmenorrhoea or dyspareunia associated with ≥ 1 of the following signs: pouch of Douglas (POD) tenderness or nodules, pain caused by cervical mobilisation, pain during POD mobilisation; intestinal symptoms alone not considered inclusion criteria. Exclusion criteria: previous surgical therapy for intestinal endometriosis and previous use of medical therapy for endometriosis

Study design: prospective, observational; consecutive enrolment

Patient characteristics and setting

Clinical presentation: symptoms - see Inclusion criteria

Age: mean 35.8 ± 4.4 years, range 28 to 48 years

Number enrolled: 37 women

Number available for analysis: 37 women

Setting: University Hospital, Santa Casa Medical School, referral centre for endometriosis

Place of study: São Paulo, Brazil

Period of study: January 2004 to January 2005

Language: English

Index tests

Index test: DCBE; TRUS (TrEUS)

Description of positive case definition by index test as reported: DCBE - features suggestive for rectosigmoid endometriosis included spiculation, circumferential narrowing of the bowel and mass effect; unclear whether prespecified criteria or description of findings; TrEUS - criteria for rectosigmoid endometriosis referenced to primary source

Examiners: both tests performed in all patients in a non-randomised sequence, by 2 blinded examiners: DCBE - performed by a single operator under supervision of a radiologist technician; images

Ribeiro 2008a (Continued)

were then reviewed by a skilled radiologist; **TrEUS** - performed by a senior echographer, single operator; unclear whether examiners were blinded to clinical data

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: rectosigmoid endometriosis

Prevalence of target condition in the sample: DIE 37/37 (100%); rectosigmoid endometriosis 27/37 (72.9%)

Reference standard: laparoscopy 37/37 (100%) + histopathology

Description of positive case definition by reference test as reported: DIE considered when fibrotic nodules with macroscopic appearance > 5 mm were present; rectosigmoid endometriosis confirmed with histopathological examination of resected specimens, and considered when histological infiltration of any layer, from serosa to mucosa; endometriosis diagnosed when ectopic glands or stroma was found; surgical procedure described in detail

Examiners: numbers or level of expertise of surgeons not reported; all biopsies studied by the same pathologist; level of expertise not reported; not blinded to results of the index tests

Flow and timing

Time interval between index test and reference standard: 1 to 3 months

Withdrawals: none reported

Comparative

Key conclusions by the authors

Our data, although limited by sample size, confirmed that DCBE is an effective method in the diagnosis of intestinal DIE, and that DCBE can predict bowel infiltration caused by endometriosis. Tr EUS proved highly effective in the diagnosis of intestinal endometriosis

Conflict of interests

Not reported

Notes

Reported accuracy estimates for DCBE and Tr EUS for diagnosis of bowel endometriosis incorrect (wrong formulas used in the published paper)

Data on comparison between index tests presented by study authors – not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?	Yes		
--	-----	--	--

Did the study avoid inappropriate exclusions?	No		
---	----	--	--

Was a 'two-gate' design avoided?	Yes		
----------------------------------	-----	--	--

High

Low

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge	Yes		
---	-----	--	--

Ribeiro 2008a *(Continued)*
of the results of the reference
standard?

Did the study provide a clear pre-specified definition of what was considered to be a “positive” result of index test?	Yes
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes
Low	

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	No
High	
Low	

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Low	

Said 2014
Study characteristics

Patient sampling	Primary objectives: to evaluate different ultrasonic signs (soft markers) for prediction of endometriosis in women of reproductive age with symptoms suggestive of endometriosis but with normal ovarian size and no evidence of ovarian cyst
------------------	--

Said 2014 (Continued)

Study population: women with any symptoms suggestive of endometriosis who were booked for laparoscopy

Selection criteria: Inclusion criteria: reproductive age; pain in the lower abdomen or pelvis for ≥ 6 months; infertility; regular menstrual cycle; no medications for infertility or pelvic pain treatment in the preceding 3 months; availability of complete past medical, social, obstetrical and gynaecological history; normal size ovary on TVS. Exclusion criteria: virginity, pregnancy, ovarian cyst of any type on TVS, genital malformation that made examination or TVS impossible, history of gynaecological cancer or previous abdominal or pelvic surgery, premature ovarian failure, large uterine masses

Study design: prospective, observational; consecutive enrolment

Patient characteristics and setting

Clinical presentation: dysmenorrhoea 96/142, dyspareunia 72/142, dyschezia 33/142, non-cyclical chronic pelvic pain 28/142, infertility 37/142, dysuria 5/142

Age: median 29 years, range 19 to 46 years

Number enrolled: 142 women

Number available for analysis: 125 women

Setting: University Hospital, El-Shatby Maternity Hospital, Alexandria University

Place of study: Alexandria University, Egypt

Period of study: not specified

Language: English

Index tests

Index test: TVUS (TVS)

Description of positive case definition by index test as reported: systematic assessment of pelvic structures with focus on sift markers for endometriosis as follows: ovaries not at same level, adhesions between ovaries and uterus, fixation of ovaries to iliac vessels (fixed non-sliding ovaries across surrounding structures), peritoneal cysts in POD or vesicouterine pouch, POD free fluid, POD obliteration (no sliding between serosa on the posterior surface of the cervix/uterus and bowel when the uterus was gently mobilised by a combination of pressure on the cervix with the ultrasound probe alternating with pressure on the fundus from the examiner's free hand; examination protocol described in detail and included a tenderness-guided approach

Examiners: TVS performed by an experienced sonographer; unclear whether blinded to clinical data

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: pelvic endometriosis

Prevalence of target condition in the sample: pelvic endometriosis 68/125 (64.4%)

Reference standard: laparoscopy 125/125 (100%) + histopathology

Description of positive case definition by reference test as reported: documentation of endometriosis by laparoscopy done by visual diagnosis of endometriotic spots, electrocoagulation test (done by applying bipolar current to suspected lesions - formation of black spots due to haemosiderin indicates biopsy taking) or biopsy from the lesion in uncertain cases; surgical procedure not described

Examiners: no data provided

Flow and timing

Time interval between index test and reference standard: 0 to 5 days

Withdrawals: 17 women excluded because of the presence of hard markers for endometriosis - an exclusion criterion

Comparative

Said 2014 (Continued)

Key conclusions by the authors	In conclusion, TVS appears to be a useful imaging method for prediction of endometriosis. However, good training, skills and passion are prerequisites for the sonographer carrying out the examination. Inclusion of TVS-based soft markers in women with symptoms suggestive of endometriosis improves our ability to predict or exclude the presence of endometriosis
Conflict of interests	Study authors declared no conflict of interest
Notes	<p>Reported accuracy estimates for TVS for diagnosis of pelvic endometriosis in women with normal ovarian size confirmed as accurate</p> <p>Association between soft markers and laparoscopic findings as well as accuracy estimates for each soft marker presented - not included in the review</p> <p>Study authors present estimates of improved diagnostic accuracy with incorporation of the 6 'positive soft markers' - we were not able to reproduce these estimates and have not presented these data</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be	Yes		

Said 2014 (Continued)
available when the test is
used in practice?

	Low	Low
DOMAIN 3: Reference Standard		
Is the reference stan- dards likely to correctly classify the target condi- tion?	Yes	
Were the reference stan- dard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference stan- dard?	Yes	
Did all patients receive the same reference stan- dard?	Yes	
Were all patients includ- ed in the analysis?	Yes	
	Low	

Savelli 2011

Study characteristics

Patient sampling	<p>Primary objectives: to compare the diagnostic accuracy of transvaginal sonography (TVS) and double-contrast barium enema (DCBE) for preoperative detection of deep infiltrating endometriosis (DIE) of the posterior compartment</p> <p>Study population: patients with results of pelvic examination or symptoms suggestive of DIE of the posterior compartment</p> <p>Selection criteria: Inclusion criteria: symptoms or examination findings indicative of DIE of the posterior compartment</p> <p>Study design: prospective, observational; consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: infertility 30/69, dysmenorrhoea 64/69, dyspareunia 59/69, dyschezia 45/69; nulliparous 49/69, previous surgery for endometriosis 18/69, oestrogen-progestin therapy before surgery 22/69</p> <p>Age: median 33.6 ± 5.9 years</p>

Savelli 2011 (Continued)

Number enrolled: 94 women

Number available for analysis: 69 women

Setting: university hospital tertiary care referral, S. Orsola-Malpighi Hospital

Place of study: Bologna, Italy

Period of study: January 2004 to December 2007

Language: English

Index tests

Index test: TVUS (TVS); DCBE

Description of positive case definition by index test as reported: TVS - features of endometriosis including presence of hypoechoic linear thickening and/or hypoechoic irregular-shaped nodules with hyperechoic rim and scarcely vascularised at power Doppler on assessed pelvic structures and/or tender fixed nodule with pain evoked by exerting gentle pressure with the vaginal probe; the following pelvic structures were assessed: uterus, ovaries, rectosigmoid colon, pouch of Douglas, uterosacral ligaments, rectovaginal septum and posterior vaginal wall; unclear whether prespecified criteria or description of findings; DCBE - diagnostic criteria referenced to a primary source and described

Examiners: both DCBE and TVS performed by 2 groups of physicians specialising in endometriosis with training and expertise in gynaecological imaging studies, who were aware of each patient's history, symptoms and pelvic examination but were blinded to the results of other index tests

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: posterior DIE, rectosigmoid endometriosis

Prevalence of target condition in the sample: posterior DIE 67/69 (97%), rectosigmoid endometriosis 56/69 (81.2%)

Reference standard: laparoscopy 69/69 (100%) + histopathology

Description of positive case definition by reference test as reported: surgeon scored the stage of pelvic endometriosis by eye subjectively (rAFS classification); confirmation of DIE based on the presence of endometrial glands and stroma together with fibrosis and smooth muscle cell hyperplasia and hypertrophy; surgical procedure described in detail

Examiners: laparoscopy performed by a skilled gynaecological surgeon specialising in endometriosis (single operator), who was aware of TVS and DCBE findings; data on numbers or level of expertise of pathologists not provided

Flow and timing

Time interval between index test and reference standard: within 1 month

Withdrawals: 25 (27%) enrolled participants excluded because of refusal to undergo surgical intervention (n = 15), refusal to undergo DCBE (n = 7) or incomplete TVS examination because this was judged to be painful by the woman, who asked that the examination be stopped (n = 3)

Comparative

Key conclusions by the authors

TVS has much higher sensitivity than DCBE in detecting the presence of posterior DIE and should thus be regarded as the imaging modality of choice when the disease is clinically suspected, reserving DCBE for cases of signs and symptoms strongly suggestive of the presence of bowel DIE in the upper part of the sigmoid, which is difficult to visualise on TVS

Conflict of interests

Not reported

Notes

Reported accuracy estimates for TVS and DCBE for diagnosis of posterior DIE confirmed as accurate; reported statistic for bowel DIE incorrect (cases with non-bowel endometriosis excluded from 2 × 2 tables in the published paper)

Savelli 2011 (Continued)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	No		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	No		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
		High	Low

Savelli 2011 (Continued)

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? No

High

Scarella 2013
Study characteristics

Patient sampling **Primary objectives:** to evaluate the diagnostic potential of transvaginal ultrasound with bowel preparation (USTV-PI) for detection of deep endometriosis

Study population: women with chronic pelvic pain and/or suspected endometriosis

Selection criteria: exclusion criteria: postmenopausal patients, patients with previous surgery of colon/sigmoid, patients with known causes of pelvic pain

Study design: multi-centre, cross-sectional; consecutive enrolment

Patient characteristics and setting **Clinical presentation:** infertility 29/57, moderate to severe pelvic pain 50/57, dyspareunia 30/57; nulliparous 30/57

Age: women of reproductive age, age range or mean not specified

Number enrolled: 100 women

Number available for analysis: 57 women

Setting: 2 university hospitals: Institute of Maternal and Child Research, University of Chile; Center for Human Reproduction, Valparaíso University

Place of study: Santiago and Valparaíso, Chile

Period of study: September 2011 to September 2012

Language: Spanish

Index tests **Index test:** TVUS (USTV-PI, with bowel preparation)

Description of positive case definition by index test as reported: each exam interpreted in real time and documented in photographs; deep endometriosis defined as presence of ≥ 1 thick nodular hypoechoic lesion in the following areas: bladder, vesicouterine, ureteral meatus, uterus, ovaries, POD, retrocervical space, USL, rectovaginal septum, vaginal fornix, rectosigmoid

Examiners: all examinations performed by a single experienced examiner; blinding to clinical data not reported

Interobserver variability: not provided

Scarella 2013 (Continued)

Target condition and reference standard(s)	Target condition: deep pelvic endometriosis (DE) - overall and separate anatomical sites; ovarian endometriosis Prevalence of target condition in the sample: deep pelvic endometriosis 35/57 (61%), ovarian endometriosis 31/57 (54%) Reference standard: laparoscopy, laparotomy (numbers for each procedure not specified) + histopathology Description of positive case definition by reference test as reported: endometriosis biopsy report and/or direct visualisation by surgeon considered as gold standard - criteria not specified; procedure not described Examiners: numbers or level of expertise of surgeons or pathologists not reported; examiners blinded to ultrasound results		
Flow and timing	Time interval between index test and reference standard: within 6 months (communication with study authors) Withdrawals: 43 (43%) of enrolled patients did not undergo surgery: pending surgery (n = 22), postponed surgery to undergo assisted reproduction technique (n = 21)		
Comparative			
Key conclusions by the authors	USTV is an appropriate test for evaluation of deep endometriosis; confirms the importance of this technique for defining a surgical strategy and appropriate counselling		
Conflict of interests	Not reported		
Notes	Reported accuracy estimates for transvaginal ultrasound for diagnosis of ovarian and deep pelvic endometriosis as well as for specific sites (USL and retrocervical) confirmed as accurate Accuracy parameters for diagnosis of bowel involvement reported as 100% - no clear raw data available; therefore not presented in this review		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was	Yes		

Scarella 2013 (Continued)

considered to be a “positive” result of index test?

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice? Unclear

Unclear

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? No

High

Stabile 2013
Study characteristics

Patient sampling

Primary objectives: to evaluate the accuracy of water enema computed tomography (CT) for prediction of location of endometriosis in patients for whom magnetic resonance imaging (MRI) is contraindicated, focusing on rectosigmoid lesions and using laparoscopic and histological data as the reference standard

Study population: women suspected to have deep pelvic endometriosis (DPE) and bowel endometriosis based on history and findings of physical examination

Selection criteria: Inclusion criteria: clinical symptoms suggestive of bowel endometriosis. Exclusion criteria: difficult and painful rectosigmoid endoscopy due to anomalous narrowing of bowel lumen caused by extrinsic compression; video laparoscopy within 4 weeks of CT examination; ab-

Stabile 2013 (Continued)

	<p>solute and relative contraindications to MRI (MR-incompatible metallic implants or known claustrophobia)</p> <p>Study design: prospective, observational; consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: chronic pelvic pain, dysmenorrhoea, dyspareunia, infertility and gastrointestinal disorders suggestive of bowel involvement (rectal pain co-incident with menses and cramping abdominal pain before or during passage of stools, defecation disorders without signs of bowel obstruction); no patients had a previous history of major abdominal surgery; previous appendectomy 4/37; previous surgery for endometriosis 6/33</p> <p>Age: mean 31.5 ± 3.4 years, range 24 to 39 years</p> <p>Number enrolled: 37 women</p> <p>Number available for analysis: 33 women</p> <p>Setting: University Hospital, University of Bari Medical School</p> <p>Place of study: Bari, Italy</p> <p>Period of study: May 2009 to December 2010</p> <p>Language: English</p>
Index tests	<p>Index test: MDCT-e (water enema CT)</p> <p>Description of positive case definition by index test as reported: all images analysed on a dedicated workstation (HP XW 8600, Minnetonka, MN, USA), with image reconstruction software (Vitrea FX 2.1, Vital Images, Minneapolis, MN, USA); diagnostic criteria for rectosigmoid endometriosis referred to a primary source and described</p> <p>Examiners: 2 radiologists with 15 years' and 5 years' experience in abdominal imaging, who were blinded to clinical data and to other results</p> <p>Interobserver variability: almost perfect agreement was found between the 2 readers (k = 0.84)</p>
Target condition and reference standard(s)	<p>Target condition: rectosigmoid endometriosis</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 33/33 (100%), DPE 26/33 (78.8%), rectosigmoid endometriosis 23/33 (69%)</p> <p>Reference standard: laparoscopy 33/33 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: diagnosis of rectosigmoid endometriosis based on presence of ectopic endometrial and stromal tissue penetrating at least into the serosal layer of the bowel wall; surgical procedure described</p> <p>Examiners: surgeon with 15 years' experience in abdominal video laparoscopy (single operator); data on numbers or level of expertise of pathologists not provided; unclear whether blinded to results of the index test</p>
Flow and timing	<p>Time interval between index test and reference standard: within 4 weeks</p> <p>Withdrawals: no withdrawals reported</p>
Comparative	
Key conclusions by the authors	<p>Water enema CT can play a role in the diagnosis of bowel endometriosis and represents another accurate potential tool for video laparoscopic approaches, especially in patients for whom MRI is contraindicated</p>
Conflict of interests	Not reported

Stabile 2013 (Continued)

Notes Reported accuracy estimates for water enema CT for diagnosis of rectosigmoid endometriosis confirmed as accurate

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Yes		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Stabile 2013 (Continued)

	Unclear	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Stratton 2003

Study characteristics	
Patient sampling	<p>Primary objectives: to evaluate the utility of fat-suppressed magnetic resonance imaging (MRI) in the diagnosis of endometriosis</p> <p>Study population: women 18 to 45 years of age with pelvic pain, who were otherwise in good health, were evaluated to exclude other causes of pain (from a cohort of women recruited for a randomised, double-blind, placebo-controlled study of surgical excision followed by innovative medical treatment for endometriosis)</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational; unclear whether consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: pelvic pain (menstrual, coital and non-menstrual pelvic pain) confirmed by standardised questionnaire using a visual analogue scale; none treated for endometriosis in the past 6 months nor had taken hormonal medication in the past 3 months; prior surgical diagnosis of endometriosis 38/58</p> <p>Age: range 20 to 44 years</p> <p>Number enrolled: 58 women</p> <p>Number available for analysis: 46 women</p> <p>Setting: university hospitals, Warren G. Magnusen Clinical Center, National Institutes of Health, Georgetown University Medical Center</p> <p>Place of study: Bethesda, MD, Washington, DC, USA</p> <p>Period of study: January 1999 to November 2000</p> <p>Language: English</p>
Index tests	<p>Index test: <u>MRI</u> (T1/T2-w + fat-suppressed + Gd)</p> <p>Description of positive case definition by index test as reported: lesions characterised by signal intensity (high, low or isodense to adjacent muscle) on unenhanced T1-w and T2-w sequences and whether they showed enhancement with gadolinium contrast. An attempt was made to diagnose all implants, including superficial ones. No attempt was made to diagnose adhesions. For an individual patient, the diagnosis of endometriosis by MRI was considered to be positive when it correlated with ≥ 1 biopsy-proven lesion</p>

Stratton 2003 (Continued)

Examiners: 2 experienced, board-certified radiologists analysed preoperative magnetic resonance images and recorded a consensus reading of the extent and location of possible endometriosis. Radiologists were aware of the clinical possibility of deep endometriosis in all participants but did not know the results of surgery, pelvic ultrasound, history, physical exam findings or histopathology

Interobserver variability: not provided

Target condition and reference standard(s)	<p>Target condition: pelvic endometriosis</p> <p>Prevalence of target condition in the sample: pelvic endometriosis 44/48 (91.6%) women, ASRM I to II 29 (66%) women, ASRM III to IV 15 (34%) women</p> <p>Reference standard: laparoscopy 48/48 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: peritoneal lesions categorised by colour, depth and width. Lesion colour categorised as follows: [1] blue, black or brown; [2] red or clear; [3] white or yellow; or [4] a mixture of the other categories. Lesions categorised as endometriomas, peritoneal defects, deep lesions, superficial lesions and small lesions. Endometriosis measuring 1 cm below the surface considered to be deep. All lesions excised and examined histologically for confirmation of endometriosis with glands or stroma. Hemosiderin-laden macrophages not considered sufficient for the diagnosis of endometriosis; surgical procedure described in detail</p> <p>Examiners: surgical team included ≥ 1 of 2 authors; level of expertise not reported. Surgeons and radiologists unaware of each other's findings</p>
Flow and timing	<p>Time interval between index test and reference standard: within 1 month (personal communication with study authors)</p> <p>Withdrawals: 12 women (21%) for the following reasons: 3 dropped out, 4 had other causes of pelvic pain (pelvic inflammatory disease, fibroids, musculoskeletal pain) and 5 did not meet entry criteria (morbid obesity, bipolar disorder or major depression or did not undergo MRI)</p>
Comparative	
Key conclusions by the authors	Although MRI identified fewer areas of endometriosis than were seen at surgery, it suggested endometriosis in 75% of those with at least mild disease. Only 67% of lesions identified at surgery contained histological evidence of endometriosis
Conflict of interests	Not stated. Study supported by the Intramural Program, National Institute of Child Health and Human Development, Bethesda, Maryland
Notes	<p>Reported accuracy estimates for MRI for diagnosis of pelvic endometriosis confirmed as accurate</p> <p>Data for different stages of endometriosis and for different size of lesions reported per number of lesions; does not allow construction of 2×2 tables - not presented in this review</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		

Stratton 2003 (Continued)

Was a 'two-gate' design avoided? Yes

Unclear

Low

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test? Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice? Yes

Low

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Stratton 2003 (Continued)

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? No

High

Sugimura 1993
Study characteristics

Patient sampling **Primary objectives:** to analyse the value of magnetic resonance imaging (MRI) in detection and characterisation of pelvic endometriosis; to assess the usefulness of fat-saturated MRI for detection of endometrial cysts, with laparoscopy or laparotomy as the standard reference

Study population: women with clinically suspected endometriosis

Selection criteria: not specified

Study design: prospective, observational; consecutive enrolment

Patient characteristics and setting

Clinical presentation: not specified

Age: mean 36 years, range 24 to 48 years

Number enrolled: 35 women

Number available for analysis: 35 women

Setting: university hospital, Shimane Medical University

Place of study: Izumo, Japan

Period of study: March 1991 to August 1992

Language: English

Index tests

Index test: MRI (T1/T2-w)

Description of positive case definition by index test as reported: assessed sites included surface of the uterus, adnexa, POD, peritoneum, ovaries; recorded details included location of lesion, size, shape; thickness, regularity and signal intensity of lesion margins; distinctness of interface of the lesion with adjacent organs; and appearance of the lesion. Criteria provided only for ovarian endometrioma and referenced to a primary source

Examiners: MRI images prospectively read by 2 study authors who were aware that patients had a clinical history of suspected endometriosis; level of expertise not reported

Interobserver variability: not provided

Target condition and reference standard(s)

Target condition: pelvic endometriosis

Prevalence of target condition in the sample: pelvic endometriosis 26/35 (74.3%)

Reference standard: laparoscopy 13/35 (37%), laparotomy 22/35 (63%) + histopathology

Description of positive case definition by reference test as reported: diagnostic criteria not mentioned; surgical procedure not described

Sugimura 1993 (Continued)

Examiners: 'laparoscopy and laparotomy procedure reports and photographs and histologic slides (when available) were reviewed by 2 gynaecologists from our university' - additional information not provided; unclear whether blinded to results of the index test

Flow and timing	Time interval between index test and reference standard: within 2 weeks Withdrawals: no withdrawals reported
Comparative	
Key conclusions by the authors	Diagnostic accuracy improved with addition of fat-saturated images, so their use together with conventional images is recommended in assessment of endometriosis
Conflict of interests	Not reported
Notes	Reported accuracy estimates for conventional MRI for diagnosis of pelvic endometriosis confirmed as accurate Data for endometrioma reported separately for large and small endometriomas; this does not allow construction of 2 × 2 tables - not presented in this review Likely overlap with data for fat-saturated MRI for another larger study from the same group - Okada 1995 (study period August 1991 to December 1993) - not able to clarify with study authors; therefore these data have been removed from the index study

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	No		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	No		
Were the same clinical data available when the index test results were interpreted as would be	Yes		

Sugimura 1993 *(Continued)*

available when the test is used in practice?

	High	Low
DOMAIN 3: Reference Standard		
Is the reference standards likely to correctly classify the target condition?	Yes	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	
	Unclear	Low
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Takeuchi 2005
Study characteristics

Patient sampling	<p>Primary objectives: to evaluate the usefulness of the MRI jelly method as a preoperative diagnostic means for patients with rectovaginal endometriosis</p> <p>Study population: women scheduled to undergo laparoscopy for suspected rectovaginal endometriosis based on clinical symptoms, rectal/pelvic examination findings and preoperative sonographic examination results</p> <p>Selection criteria: not specified</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: dysmenorrhoea 31/31, dyspareunia 10/31, chronic pelvic pain 7/31; sonography suggestive for endometrioma 25/31; none had a history of previous pelvic surgery, and none had received hormonal therapy within 6 months preceding the study</p> <p>Age: mean 32.1 ± 4.2 years</p> <p>Number enrolled: 31 women</p> <p>Number available for analysis: 31 women</p> <p>Setting: university hospital, Juntendo University School of Medicine</p>

Takeuchi 2005 (Continued)

Place of study: Tokyo, Japan

Period of study: January 2001 to July 2002

Language: English

Index tests	<p>Index test: <u>MRI</u> (T1/T2-w + fat-suppressed, jelly method)</p> <p>Description of positive case definition by index test as reported: cul-de-sac (CDS) obliteration defined as unnatural flexure of the posterior wall of the uterine cervix or thickening or tension of rectal folds on T2-w images. Areas of low intensity in CDS defined as deep lesions and classified as thick or nodular lesions (posterior wall of the uterine cervix, posterior vaginal vault and anterior wall of the rectum)</p> <p>Examiners: MRI images read preoperatively by 1 radiologist who was blinded to clinical findings; level of expertise not reported</p> <p>Interobserver variability: not provided</p>		
Target condition and reference standard(s)	<p>Target condition: posterior DIE (posterior deep pelvic endometriosis), POD obliteration (CDSO)</p> <p>Prevalence of target condition in the sample: posterior deep pelvic endometriosis 17/31 (55%), CDSO 22/31 (71%)</p> <p>Reference standard: laparoscopy 31/31 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: presence or type of CDSO as well as deep lesions found after opening of the CDSO (deep lesions adhering to posterior wall of the uterine cervix, posterior vaginal vault or anterior wall of the rectum) evaluated according to ASRM classification; cul-de-sac diagnosed as normal when the bulge of the posterior vaginal vault was between both USLs, as partial cul-de-sac obliteration (PCDSO) when only part of the bulge of the posterior fornix was seen, and as complete cul-de-sac obliteration (CCDSO) when the posterior vaginal vault could not be seen at all; surgical procedure described in detail</p> <p>Examiners: numbers or level of expertise of surgeons or pathologists not reported; surgeon blinded to MRI findings</p>		
Flow and timing	<p>Time interval between index test and reference standard: not reported, but statement "before surgery jelly for ultrasonography was injected into the vagina and rectum for MRI" suggests that imaging was performed immediately before surgery</p> <p>Withdrawals: none reported</p>		
Comparative			
Key conclusions by the authors	The condition of the cul-de-sac could be imaged clearly via the MRI jelly method. Not only rectovaginal endometriosis presenting with deep lesions, but also complete cul-de-sac obliteration alone, could be diagnosed preoperatively at a high rate		
Conflict of interests	Not reported		
Notes	<p>Reported accuracy estimated for MRI jelly method for diagnosis of obliterated CDS and deep endometriosis confirmed as accurate</p> <p>Previous US suggestive for endometriosis reported in 81% of enrolled participants; however, operator of index test blinded to these data; no participants had previous pelvic surgery</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns

Takeuchi 2005 (Continued)

DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?

No

Did the study avoid inappropriate exclusions?

Unclear

Was a 'two-gate' design avoided?

Yes

High

Low

DOMAIN 2: Index Test Any test

Were the index test results interpreted without knowledge of the results of the reference standard?

Yes

Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?

Yes

Was the index test performed by a single operator or interpreted by consensus in a joint session?

Yes

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?

Yes

Low

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?

Yes

Were the reference standard results interpreted without knowledge of the results of the index tests?

Yes

Low

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?

Unclear

Takeuchi 2005 (Continued)

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Unclear

Thomeer 2014
Study characteristics

Patient sampling

Primary objectives: to explore whether an optimised 3.0Tesla MRI protocol for endometriosis may be more sensitive than laparoscopic exploration for detecting the disease in a clinical setting

Study population: patients with clinical suspicion of endometriosis scheduled to undergo laparoscopy

Selection criteria: exclusion criteria: use of contraceptives or hormonal suppressive medication, contraindication to MRI (pacemaker, different metallic bodies, claustrophobia), age younger than 18, post-menopausal status

Study design: prospective, observational; consecutive enrolment

Patient characteristics and setting

Clinical presentation: pain, subfertility and other symptoms suggestive of endometriosis (not specified)

Age: median 25 years, range 18 to 39 years

Number enrolled: 40 women

Number available for analysis: 40 women

Setting: university hospital, Erasmus Medical Centre, Rotterdam University

Place of study: Rotterdam, The Netherlands

Period of study: November 2010 to December 2012

Language: English

Index tests

Index test: MRI 3.0T

Description of positive case definition by index test as reported: diagnosis of an endometrioma based on shading on T2-weighted images and hyperintensity on T1-weighted images. Fibrotic-like tissue on T2-weighted images stated as deep endometriosis. Focal T1-weighted hyperintense foci without T2-weighted abnormalities considered as superficial endometriotic lesions. POD obliteration based on visibility of adhesions between uterus and bowel loops; referenced to primary source

Examiners: 2 experienced radiologists (blinded), with 13 years' and 12 years' experience in abdominal MRI, analysed independently and blindly data on a PACS workstation. They had no information regarding clinical data; disagreements about image interpretation were sorted by consensus

Interobserver variability: perfect per-patient interobserver agreement ($k = 1$); substantial per-lesion interobserver agreement ($k = 0.65$)

Target condition and reference standard(s)

Target condition: pelvic endometriosis, POD obliteration

Prevalence of target condition in the sample: pelvic endometriosis 37/40 (92.5%), rASRM I to II 20/37 (54%), rASRM III to IV 17/37 (46%), POD obliteration 10/40 (25%)

Thomeer 2014 (Continued)

Reference standard: laparoscopy 40/40 (100%)

Description of positive case definition by reference test as reported: location and size of each peritoneal lesion and endometrioma found during laparoscopy recorded with digital video and reviewed by gynaecologists (blinded); staging - according to rASRM criteria; lesion classified as superficial if it was stated as no deeper than 4 mm below the peritoneal surface; obliteration of the cul-de-sac mentioned if the cul-de-sac was not accessible from peritoneal space because of the presence of adhesions

Examiners: operative videos reviewed by 2 gynaecologists with extensive experience with laparoscopy and detecting endometriosis; interobserver agreement with consensus reading performed; readers blinded to MRI findings; no data provided on the team performing surgery

Flow and timing	<p>Time interval between index test and reference standard: within 2 months</p> <p>Withdrawals: none reported</p>
Comparative	
Key conclusions by the authors	In conclusion, with an optimised protocol, MRI seems reliable for detecting all patients with endometriosis higher than stage I. Additional studies in larger patient populations and in patients with low suspicion of the disease are needed to confirm our findings
Conflict of interests	Study authors declared no conflict of interest
Notes	<p>Reported accuracy estimates for MRI for diagnosis of pelvic endometriosis and POD obliteration confirmed as accurate</p> <p>Diagnostic estimates for different pelvic regions and for superficial vs deep lesions reported - not presented in the review because of 'lesion-type' analysis</p> <p>Study authors also present an explorative algorithm to classify patients as having low stage vs high stage of endometriosis; data insufficient for calculating accuracy estimates for each stage</p>

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of the index test?	Yes		

Thomeer 2014 *(Continued)*

ition of what was considered to be a “positive” result of index test?

Was the index test performed by a single operator or interpreted by consensus in a joint session? Yes

Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice? Yes

Low

Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Low

Low

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Ubaldi 1998
Study characteristics

Patient sampling

Primary objectives: to investigate prospectively the role of TVUS in infertile women undergoing laparoscopy for infertility workup and to analyse the predictive value of TVUS in differentiating normal from pathological pelvis

Study population: patients who had been referred for diagnostic or operative laparoscopy for infertility, chronic pelvic pain and/or adnexal masses

Ubaldi 1998 (Continued)

	<p>Selection criteria: Inclusion criteria: non-pregnant premenopausal women</p> <p>Study design: prospective, observational; non-consecutive enrolment</p>
Patient characteristics and setting	<p>Clinical presentation: infertility, chronic pelvic pain and/or adnexal masses</p> <p>Age: range 21 to 41 years</p> <p>Number enrolled: 133 women</p> <p>Number available for analysis: 133 women</p> <p>Setting: university hospital: Centre for Reproductive Medicine of the Dutch-speaking Free University of Brussels</p> <p>Place of study: Brussels, Belgium</p> <p>Period of study: February 1994 to April 1995</p> <p>Language: English</p>
Index tests	<p>Index test: <u>TVUS</u></p> <p>Description of positive case definition by index test as reported: criteria for diagnosis of ovarian endometriosis: thick walls, regular margins and homogeneous low echogenicity of fluid. Scan considered normal when the uterus had normal morphology, with no uterine contour or positional abnormalities, and adnexae were in their anatomical position, free of ovarian masses, hydrosalpinges or other pathologies</p> <p>Examiners: all scans performed by 2 physicians, each with ≥ 3 years' expertise in ultrasound scanning; physicians not told about clinical histories of patients</p> <p>Interobserver variability: presented for differentiating normal pelvis from any pelvis pathology but not specifically for diagnosis of endometrioma</p>
Target condition and reference standard(s)	<p>Target condition: ovarian endometriosis</p> <p>Prevalence of target condition in the sample: endometrioma 10/133 (7.5%)</p> <p>Reference standard: laparoscopy 133/133 (100%) + histopathology</p> <p>Description of positive case definition by reference test as reported: uterus, pelvic peritoneum and ovaries carefully observed; severity of endometriosis evaluated according to ASRM classification; surgical procedure and histological criteria not described.</p> <p>Examiners: numbers or level of expertise of surgeons or pathologists not reported; unclear whether surgeons blinded to index test findings</p>
Flow and timing	<p>Time interval between index test and reference standard: 1 day</p> <p>Withdrawals: none reported</p>
Comparative	
Key conclusions by the authors	<p>In conclusion, the present study suggests that TVUS is accurate in diagnosing pelvic pathologies except for filmy pelvic adhesions. In the initial workup of infertile women, if hysterosalpingography demonstrates patent tubes, if the patient is young and if TVUS is negative, laparoscopy can be postponed. In couples with severe male factor infertility for whom in vitro fertilisation or intracytoplasmic sperm injection is the treatment of choice, if TVUS is negative, laparoscopy may be avoided</p>
Conflict of interests	Not reported

Ubaldi 1998 (Continued)

Notes

Reported accuracy estimated for TVUS for diagnosis of endometrioma confirmed as accurate

Accuracy estimates of TVUS for overall pelvic pathology and for pelvic adhesions presented - not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	Low
DOMAIN 2: Index Test Any test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
Did the study provide a clear pre-specified definition of what was considered to be a "positive" result of index test?	Yes		
Was the index test performed by a single operator or interpreted by consensus in a joint session?	Unclear		
Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	Low

Ubaldi 1998 *(Continued)*
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low

Abbreviations:

2D: Two-dimensional.

3D: Three-dimensional.

3D-TVUS: Three-dimensional transvaginal ultrasound.

3D-US: Three-dimensional ultrasound.

BE: Barium enema.

CCDSO: Complete cul-de-sac obliteration.

CDS: Cul-de-sac.

CDSO: Cul-de-sac obliteration.

CSE: Conventional spin echo.

CT: Computed tomography.

DCBE: Double-contrast barium enema.

DIE: Deep infiltrating endometriosis.

DIPE: Deep infiltrating posterior endometriosis.

¹⁸FDG PET-CT: Fluorodeoxyglucose positron emission tomography/computed tomography.

FSE: Fast spin echo.

Gd-TIFS: Gadolinium-enhanced T1-weighted fat-suppressed.

GnRH: Gonadotropin-releasing hormone.

MDCT-e: Multi-detector computerised tomography enteroclysis.

MRI: Magnetic resonance imaging.

MRI-e: Magnetic resonance imaging enema.

MSCT: Multi-slice computed tomography.

NPV: Negative predictive value.

PACS: Picture archiving and communication system.

PCDSO: Partial cul-de-sac obliteration.

POD: Pouch of Douglas.

PPV: Positive predictive value.

RES: Rectal endoscopic sonography.

RS: Rectosigmoid.

RVS: Rectovaginal septum.

RWC-TVSL: Rectal water contrast transvaginal ultrasonography.

tg-TVUS: Tenderness-guided transvaginal ultrasound.

TIFS: T1-weighted fat-suppressed.

Tr EUS: Transrectal endoscopic ultrasonography.

TRS: Transrectal sonography.

TVS: Transvaginal sonography.

TVUS: Transvaginal ultrasonography.

TVUS-BP: Transvaginal ultrasonography with bowel preparation.

USL: Uterosacral ligaments.

USTV-PI: Transvaginal ultrasound with bowel preparation.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Abrao 2004	Insufficient diagnostic test accuracy information (unable to construct 2 × 2 tables); unable to contact study authors
Alcazar 1997	Population and outcome outside inclusion criteria (postmenopausal women included; only 'lesion-level' analysis)
Alcazar 2010	Study design and outcome outside inclusion criteria (retrospective selection of cases; only 'lesion-level' analysis)
Alcazar 2011	Study outcome outside inclusion criteria (only 'lesion-level' analysis)
Anaf 2009	Not a DTA study (focus on depth of invasion of endometriotic lesions)
Arrive 1989	Time flow not in line with inclusion criteria (no data on time interval between index test and reference standard; study authors contacted - files not available)
Ayida 1997	Target condition outside inclusion criteria (focus on any kind of pelvic pathology in infertile population)
Bahr 2006	Time flow not in line with inclusion criteria (time interval between index test and reference standard exceeded 12 months)
Bazot 2003	Population overlapped with Bazot 2009
Bazot 2004a	Population outside inclusion criteria (postmenopausal women included); population overlapped with Bazot 2009
Bazot 2004b	Population outside inclusion criteria (postmenopausal women included); population overlapped with Bazot 2009
Bazot 2007a	Population overlapped with Bazot 2009
Bazot 2007b	Population overlapped with Bazot 2009
Bazot 2011a	Study design outside inclusion criteria (retrospective selection of cases)
Bazot 2011b	Study design outside inclusion criteria (retrospective selection of cases)
Bazot 2012	Study design outside inclusion criteria (retrospective selection of cases)
Bekiesinska-Figatowska 2014	Study design outside inclusion criteria (retrospective selection of cases)
Benaceraff 2015	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions; retrospective selection of cases)
Boog 1987	Study design outside inclusion criteria (retrospective selection of cases; insufficient description of methods and population)
Božidar 2010	Population outside inclusion criteria (patients with ectopic pregnancy included)
Brazert 2001	Insufficient description of methods and population; unable to contact study authors
Busard 2010	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions)
Busard 2011	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions)

Study	Reason for exclusion
Busard 2012	Study design and outcome outside inclusion criteria (retrospective selection of cases; 'lesion-level' analysis)
Busard 2014	Not a DTA study (focus on comparison between 2 types of MRI and on interobserver and intraobserver agreement; no data on surgical diagnosis)
Carbognin 2006	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Cardoso 2009	Reference standard and outcome outside inclusion criteria (no data on surgical diagnosis; only 'lesion-level' analysis)
Chamie 2009b	Population overlapped with Chamie 2009
Chapron 1998	Study design outside inclusion criteria (retrospective selection of cases)
Chapron 2004	Study design outside inclusion criteria (retrospective selection of cases)
de Kroon 2004	Population outside inclusion criteria (postmenopausal women included)
De Souza 1995	Insufficient diagnostic test accuracy information (unable to construct 2 × 2 tables)
Delpy 2005	Insufficient diagnostic test accuracy information (unable to construct 2 × 2 tables); unable to contact study authors
Demidov 1991	Insufficient description of methods and population; unable to contact study author
Di Paola 2015	Target condition outside inclusion criteria (focus on MRI-ENZIAN score); study design outside inclusion criteria (retrospective selection of cases)
Dogan 1996	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Drobne 2014	Study design outside inclusion criteria (retrospective selection of cases)
Dumontier 2000	Study design outside inclusion criteria (retrospective selection of cases); unclear whether transrectal or transvaginal ultrasound used
Egekvist 2012	Not a DTA study (focus on interobserver variability; no reference standard for 1/3 of participants)
Exacoustos 2013	Not a DTA study (focus on description of imaging findings)
Exacoustos 2014	Reference standard outside inclusion criteria (only women with positive index test underwent surgery)
Faccioli 2008	Study design and outcome outside inclusion criteria (retrospective selection of cases; only 'lesion-level' analysis)
Faccioli 2010	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Falco 1995	Study design outside inclusion criteria (retrospective selection of cases)
Fiaschetti 2012	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Fratelli 2013	Study design outside inclusion criteria (retrospective selection of cases)
Friedman 1985	Study design outside inclusion criteria (retrospective selection of cases)

Study	Reason for exclusion
Gauche Cazalis 2012	Study design outside inclusion criteria (retrospective selection of cases)
Gordon 1982	Study design outside inclusion criteria (retrospective selection of cases)
Griffiths 2008	Insufficient description of methods and population; unable to contact study authors
Guerriero 1995	Outcome outside inclusion criteria (data reported for number of lesions, not for number of patients)
Guerriero 1997	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Guerriero 1998	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Guerriero 2009	Study design outside inclusion criteria (retrospective selection of cases)
Guerriero 2010	Target condition outside inclusion criteria (focus on diagnosis of pelvic adhesions in women with suspected endometrioma)
Hauth 2004	Insufficient description of study methods and population (unclear time test to surgery, patient selection process, withdrawals); unable to clarify with study authors
Hensen 2009	Study design outside inclusion criteria (retrospective selection of cases)
Holland 2013a	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Holland 2013b	Not a DTA study (focus on interobserver variability)
Hudelist 2009a	Index test not in line with inclusion criteria (data reported for imaging test combined with examination; no separate data for imaging test)
Hudelist 2009b	Not a DTA study (focus on depth of invasion of endometriotic lesions)
Iosca 2013	Study design outside inclusion criteria (retrospective selection of cases)
Jain 1993	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Jarlot 2008	Study design outside inclusion criteria (retrospective selection of cases)
Jeong 2013	Study design outside inclusion criteria (retrospective selection of cases)
Jermy 2001	Outcome not outside inclusion criteria (only 'lesion-level' analysis)
Johnson 1994	Study design outside inclusion criteria (retrospective selection of cases)
Jung 2010	Study design outside inclusion criteria (retrospective selection of cases)
Khan 2013	Insufficient description of study methods and population (unclear age group of participants, recruitment process, time test to surgery); unable to clarify with study authors
Kikuchi 2009	Insufficient diagnostic accuracy information (separate diagnostic estimates for various radiological criteria; no overall estimates for the test)
Kikuchi 2014	Study design and target condition outside inclusion criteria (retrospective selection of cases; POD obliteration not specific for endometriosis)

Study	Reason for exclusion
Kinkel 1999	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Kreuzberg 2004	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions; no reference standard for 1/3 of participants)
Kruger 2013	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Kurjak 1994	Population and outcome outside inclusion criteria (postmenopausal women included; data reported for number of lesions, not for number of patients)
Li 2012	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions)
Li 2014	Study design, population and outcome outside inclusion criteria (retrospective selection of cases; postmenopausal women included; only 'lesion-level' analysis)
Macario 2012	Study design and outcome outside inclusion criteria (retrospective selection of cases; only 'lesion-level' analysis); focus on test performance by sign and interobserver-intraobserver performance
Mais 1993	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Mathlouthi 2011	Target condition outside inclusion criteria (benign vs malignant ovarian masses; no separate data for endometriosis)
Menada 2008b	Population overlapped with Menada 2008a
Mezzi 2011	Reference standard outside inclusion criteria (only women with positive index test underwent surgery)
Millischer 2014	Reference standard outside inclusion criteria (no data on surgical diagnosis)
Minaif 2008	Reference standard outside inclusion criteria (no data on surgical diagnosis)
Nezhat 1994	Study design outside inclusion criteria (retrospective selection of cases)
Njavro 2003	Population outside inclusion criteria (patients with ectopic pregnancy and pelvic inflammatory disease included)
Ohba 1996	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions)
Okaro 2006	Target condition outside inclusion criteria (presence vs absence of pelvic pathology; no separate data for endometriosis)
Onbas 2007	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions; retrospective selection of cases)
Outwater 1993	Study design outside inclusion criteria (retrospective selection of cases)
Pascual 2000	Population, outcome and study design outside inclusion criteria (postmenopausal women included; only 'lesion-level' analysis; retrospective selection of cases)
Pascual 2013	Not a DTA study (focus on interobserver agreement)
Patel 1999	Study design and outcome outside inclusion criteria (retrospective selection of cases; only 'lesion-level' analysis; insufficient description of population)

Study	Reason for exclusion
Pereira 2009	Not a DTA study (estimates the distance of endometriotic lesions from the anal border)
Philip 2015	Reference standard and outcome outside inclusion criteria (no data on surgical diagnosis; only 'lesion-level' analysis)
Pishvaian 2006	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions and focus on depth of invasion; retrospective selection of cases)
Preutthipan 1995	Target condition outside inclusion criteria (presence vs absence of pelvic pathology; no separate data for endometriosis)
Reid 2013b	Not a DTA study (focus on interobserver and intraobserver agreement)
Ribeiro 2008b	Population overlapped with Ribeiro 2008a
Roman 2008	Not a DTA study (focus on depth of invasion; retrospective selection of cases)
Roseau 2000	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions and focus on depth of invasion; retrospective selection of cases)
Rossi 2014	Not a DTA study (focus on depth of invasion; retrospective selection of cases)
Rousset 2014	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Roy 2009	Study design outside inclusion criteria (retrospective selection of cases)
Saba 2010	Not a DTA study (focus on interobserver and intraobserver agreement; no data on surgical diagnosis)
Saba 2011	Not a DTA study (focus on radiologists' expertise and learning curve)
Saba 2012	Reference standard outside inclusion criteria (only women with positive index test underwent surgery)
Saba 2014b	Not a DTA study (focus on interobserver variability)
Saccardi 2012	Reference standard outside inclusion criteria (only women with positive index test underwent surgery)
Scardapane 2011	Outcome outside inclusion criteria (all healthy controls excluded from analysis)
Scardapane 2013	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Scardapane 2014	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Schroder 1997	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions and correlation with intraoperative management)
Setubal 2011	Not a DTA study (description of imaging method and qualitative description of radiological appearance of endometriotic lesions)
Sherif 2015	Insufficient description of study methods and population (unclear time test to surgery, unclear whether patient-type or lesion-type analysis); unable to clarify with study authors
Sokalska 2009	Population and outcome outside inclusion criteria (postmenopausal women included; only 'lesion-level' analysis)

Study	Reason for exclusion
Stegmann 2009	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Takahashi 1994	Population appears to overlap with Sigumura 1993 and/or Okada 1995; unable to contact study authors
Takeuchi 2008	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions; retrospective selection of cases; postmenopausal women included)
Tammaa 2014	Not a DTA study (focus on learning curve of index test for diagnosis of DIE of the rectum and POD obliteration; no data on surgical diagnosis)
Tammaa 2015	Insufficient description of study methods and population (unclear time test to surgery, unclear whether patient-type or lesion-type analysis); unable to clarify with study authors
Theodoridis 2009	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Valentini 2014	Not a DTA study (qualitative description of radiological appearance of endometriotic lesions; retrospective selection of cases)
van Holsbeke 2010	Population and study design outside inclusion criteria (postmenopausal women included; retrospective selection of cases)
Vimercati 2012	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Volpi 1995	Study design outside inclusion criteria (retrospective selection of cases)
Vrachnis 2012	Target condition outside inclusion criteria (benign vs malignant ovarian masses; no separate data for endometriosis)
Weerakiet 2000	Study design, population and outcome outside inclusion criteria (retrospective selection of cases; postmenopausal women included; only 'lesion-level' analysis)
Young 2013	Study design outside inclusion criteria (retrospective selection of cases); outcome outside inclusion criteria (no separate data for imaging test, only combination imaging + clinical examination)
Zanardi 2003	Outcome outside inclusion criteria (only 'lesion-level' analysis)
Zawin 1989	Study design and outcome outside inclusion criteria (retrospective selection of cases; only 'lesion-level' analysis)
Zykin 1981	Study design outside inclusion criteria (retrospective selection of cases)

Characteristics of ongoing studies *[ordered by study ID]*

NCT01939535

Trial name or title	Preoperative Staging of Endometriosis With MRI (IDEAL) ClinicalTrials.gov Identifier: NCT01939535 Other study ID number: S54441
Target condition and reference standard(s)	Objective: to evaluate the value of MRI in preoperative stratification of endometriosis patients needing a surgical approach by gynaecologists only or a multi-disciplinary approach by gynaecologists, urologists and/or abdominal surgeons

NCT01939535 (Continued)

Primary outcome measures: odds of changed surgical approach in deep endometriosis based on preoperative MRI findings

Study design: observational, prospective

Target condition: endometriosis

Reference standard: laparoscopy

Index and comparator tests	MRI
Starting date	September 2013
Contact information	University Hospitals KULeuven, Leuven, Belgium 3000 Contact: Ingrid Fruyt +3216343781 ingrid.fruyt@uzleuven.be Contact: Linda Meersman +3216343782 linda.meersman@uzleuven.be Principal Investigator: Didier Bielen, PhD
Notes	Current status - recruiting

NCT02233621

Trial name or title	Assessment of Performance of [18F]-FES for Endometriosis Diagnosis (ENDOTEP) ClinicalTrials.gov Identifier: NCT02233621 Other study ID numbers: 49RC10_32_01-PHRC2010-02, 2011-003734-14
Target condition and reference standard(s)	Objective: to evaluate the value of MRI in preoperative stratification of endometriosis patients needing a surgical approach by gynaecologists only or a multi-disciplinary approach by gynaecologists, urologists and/or abdominal surgeons Primary outcome measures: sensitivity of PET with [18F]-FES for diagnosing endometriosis defined by the ability of this diagnostic exam to yield a positive result when endometriosis is present Study design: open label, diagnostic Target condition: endometriosis Reference standard: laparoscopy + histology
Index and comparator tests	PET with [18F]-FES (16 α -[18F]fluoro-17 β -estradiol)
Starting date	June 2012
Contact information	Nuclear Medicine Unit, University Hospital of Angers, Angers, France 49933 Contact: Olivier Couturier, PU-PH 33-(0)2-41-35-34-06 olcouturier@chu-angers.fr Contact: Céline Lefebvre-Lacoeuille, PH 33-(0)2-41-35-46-35 celefebvre@chu-angers.fr Sub-Investigator: Céline Lefebvre-Lacoeuille, PH
Notes	Current status - recruiting

NTR3738

Trial name or title	Magnetic Resonance Imaging to Diagnose Endometriosis Using Ablavar [®] as Contrast Agent: A Feasibility Study Candidate number: 14064 NTR number: NTR3738
Target condition and reference standard(s)	Hypothesis: dynamic CE-MRI using Ablavar [®] can efficiently visualise endometriosis-associated angiogenesis Primary outcome: feasibility of Ablavar [®] -enhanced MRI for detection of superficial peritoneal endometriosis using histology as the diagnostic gold standard Study design: feasibility, open-label study Target condition: peritoneal endometriosis Reference standard: laparoscopy + histology
Index and comparator tests	MRI - standard and contrast-enhanced, using gadofosveset (Ablavar) as contrast agent
Starting date	1-Feb-2013
Contact information	Dr. Andrea Romano, Maastricht University Medical Center (MUMC+)
Notes	Current status - unclear

DATA

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 TVUS pelvic	5	1222
2 TVUS ovarian	8	765
3 TVUS DIE	3	282
4 TVUS posterior DIE	7	853
5 TVUS* posterior DIE	2	248
6 TVUS USL	7	751
7 TVUS RVS	10	983
8 TVUS* RVS	1	90
9 TVUS vaginal	6	679

Test	No. of studies	No. of participants
10 TVUS POD	6	755
11 TVUS anterior DIE	2	289
12 TVUS rectosigmoid	14	1615
13 TVUS* rectosigmoid	1	202
14 TVUS bowel [ileum - rectum]	3	314
15 TRUS ovarian	1	92
16 TRUS USL	2	232
17 TRUS RVS	2	232
18 TRUS vaginal	2	232
19 TRUS rectosigmoid	4	330
20 TRUS bowel [ileum - rectum]	1	134
21 MRI pelvic	7	303
22 MRI* pelvic	2	62
23 MRI** pelvic	1	31
24 MRI ovarian	3	179
25 MRI DIE	4	212
26 MRI posterior DIE	2	54
27 MRI* posterior DIE	1	23
28 MRI USL	4	198
29 MRI* USL	1	23
30 MRI RVS	3	288
31 MRI vaginal	4	248
32 MRI* vaginal	1	23
33 MRI POD	5	154
34 MRI* POD	1	23
35 MRI anterior DIE	1	41
36 MRI rectosigmoid	6	612

Test	No. of studies	No. of participants
37 MRI+ rectosigmoid	1	23
38 MDCT-e rectosigmoid	3	389
39 MDCT-e bowel [ileum - rectum]	2	194
40 ¹⁸ FDG PET-CT pelvic	1	10
41 DCBE DIE	1	69
42 DCBE rectosigmoid	2	106
43 MRI pelvic ¹	1	35

Test 1. TVUS pelvic.

Test 2. TVUS ovarian.

Test 3. TVUS DIE.

Test 4. TVUS posterior DIE.

Test 5. TVUS+ posterior DIE.

Test 6. TVUS USL.

Test 7. TVUS RVS.

Test 8. TVUS+ RVS.

Test 9. TVUS vaginal.

Test 10. TVUS POD.

Test 11. TVUS anterior DIE.

Test 12. TVUS rectosigmoid.

Test 13. TVUS+ rectosigmoid.

Test 14. TVUS bowel [ileum - rectum].

Test 15. TRUS ovarian.

Test 16. TRUS USL.

Test 17. TRUS RVS.

Test 18. TRUS vaginal.

Test 19. TRUS rectosigmoid.

Test 20. TRUS bowel [ileum - rectum].

Test 21. MRI pelvic.

Test 22. MRI+ pelvic.

Test 23. MRI pelvic.**

Test 24. MRI ovarian.

Test 25. MRI DIE.

Test 26. MRI posterior DIE.

Test 27. MRI* posterior DIE.

Test 28. MRI USL.

Test 29. MRI* USL.

Test 30. MRI RVS.

Test 31. MRI vaginal.

Test 32. MRI* vaginal.

Test 33. MRI POD.

Test 34. MRI* POD.

Test 35. MRI anterior DIE.

Test 36. MRI rectosigmoid.

Test 37. MRI* rectosigmoid.

Test 38. MDCT-e rectosigmoid.

Test 39. MDCT-e bowel [ileum - rectum].

Test 40. ¹⁸FDG PET-CT pelvic.

Test 41. DCBE DIE.

Test 42. DCBE rectosigmoid.

Test 43. MRI pelvic¹.

ADDITIONAL TABLES

Table 1. Staging of endometriosis, rASRM classification

Peri- toneum	Endometriosis	< 1 cm	1-3 cm	> 3 cm
	Superficial	1	2	4
	Deep	2	4	6
Ovary	R Superficial	1	2	4
	Deep	4	16	20
	L Superficial	1	2	4
	Deep	4	16	20
	Posterior Cul-de-sac Obliteration	Partial Complete		
		4 40		

Table 1. Staging of endometriosis, rASRM classification (Continued)

Ovary	Adhesions	< 1/3 Enclosure	1/3-2/3 Enclosure	> 2/3 Enclosure
	R Filmy	1	2	4
	Dense	4	8	16
	L Filmy	1	2	4
	Dense	4	8	16
Tube	R Filmy	1	2	4
	Dense	4 ^a	8 ^a	16
	L Filmy	1	2	4
	Dense	4 ^a	8 ^a	16

^aIf the fimbriated end of the fallopian tube is completely enclosed, change the point assignment to 16 [American Society for Reproductive Medicine 1997](#)

Table 2. Index tests - description and common abbreviations

Test name as presented in the review	Description	Alternative names presented in the included studies
MRI tests		
MRI (magnetic resonance imaging)	<p><u>Equipment:</u> 1.5 Tesla magnet device with a parallel or phased array body or pelvic coil for signal excitation and reception</p> <p><u>Participants' preparation:</u> Fasting for 3-6 hours before the test and/or bowel preparation with oral laxatives was described by some investigators; an intravenous injection of anti-peristaltic agent at the outset of the examination to decrease bowel peristalsis; supine position. Some groups performed MRI with full bladder to correct the angle of the ante-flexed uterus; some groups described introducing of ultrasonographic gel (~ 50 to 60 mL) into the vaginal canal to distend the vaginal fornices</p> <p><u>Protocol:</u> Imaging is performed in the axial plane with or without sagittal or coronal planes. Different types of sequences allow to image the same tissue in various ways, and combinations of sequences reveal important diagnostic information about the tissue in question. The imaging parameters (section thickness, field of view (FOV), matrix size) vary between protocols. Images are documented on radiographic film and in digital files and analysed at workstation</p>	
<ul style="list-style-type: none"> MRI T1/T2-w (conventional T1-/T2-weighted) 	The protocol includes axial spin-echo or gradient echo T1-weighted (T1-w) images followed by fast spin-echo (FSE)/turbo spin-echo (TSE) images or fast relaxation fast-spin echo (FR-FSE) T2-w images	MRI; CSE (conventional spin echo)
<ul style="list-style-type: none"> MRI fat-suppressed (T1-weighted) 	Protocol includes T1-w imaging using chemical fat suppression, which aids in the differentiation of lipid and haemorrhagic pathologies. <i>Fat suppression</i> is a generic term that includes various techniques to suppress the signal from normal adipose tissue to reduce	Fat-saturated MRI

Table 2. Index tests - description and common abbreviations (Continued)

chemical shift artefact and can be achieved by various methods. This is commonly a part of the MRI protocol and is rarely used in isolation

<ul style="list-style-type: none"> MRI T1/T2-w + fat-suppressed/ Gd <p>(T1-/T2-weighted with fat-suppression contrast enhanced)</p>	<p>Protocol includes gradient echo T1 images with and without fat suppression followed by FSE or FR-FSE T2-w images before and after intravenous injection of the paramagnetic contrast agent gadolinium</p>	<p>MRI;</p> <p>CSE/TIFS (conventional spin echo in combination with T1-w fat-suppressed)</p> <p>CSE/TIFS/Gd-TIFS (conventional spin echo in combination with T1-w fat-suppressed and gadolinium-enhanced TIFS)</p>
<ul style="list-style-type: none"> MRI 'jelly method' 	<p>Protocol involves pretreatment of participants for MRI by simultaneous injection of ultrasonographic gel into the vagina (~ 50 mL) and into the rectum (150 mL gel 50% diluted with water). Another technique evolves introduction of 300-400 mL of diluted ultrasonographic gel (1:8 dilution) for rectosigmoid distension without use of intravaginal gel</p>	<p>MRI-e (magnetic resonance enema)</p>
<p>3D-MRI (3-dimensional MRI)</p>	<p>Protocol includes 3D coronal single-slab (containing all the slices) MRI, entitled 'CUBE' with FSE T2-w images. The technique involves using variable flip angle refocusing, auto-calibrating, 2D accelerated parallel imaging and nonlinear view ordering to produce high-resolution volumetric image data sets and to reduce imaging time by using multi-planar reformations</p>	
<p>3.0T MRI</p>	<p><u>Equipment:</u> 3.0Tesla Magnetom system with a multi-channel phased-array surface body-coil</p> <p><u>Participants' preparation:</u> Fasting for 3 hours before the test was reported by some but not all studies; intravenous injection of anti-peristaltic agent at the outset of the examination to decrease bowel peristalsis; administration of a negative super-paramagnetic oral contrast agent to reduce signal intensity of the bowels. Examination with the full bladder in a 'feet first' supine position</p> <p><u>Protocol:</u> combination of all or some of the following sequences: T-w FSE, 2D-T2-w FR-FSE/FSE, 3D-T2-w FR-FSE CUBE, 3D-T1-w fat-suppressed and/or LAVA-flex (liver imaging with volume acceleration-flexible) sequences. MRI images are acquired according to multiple scan planes, in particular axial, coronal and sagittal planes of the pelvis and sacral para-coronal plane. Contrast agent (gadolinium) is administered in selected cases. Total acquisition time ~ 20 min without or 30-40 min with contrast injection</p>	

Ultrasound tests

<p>TVUS</p> <p>(transvaginal ultrasonography)</p>	<p><u>Equipment:</u> any of the commercially available ultrasound machines equipped with a wide-band high-resolution vaginal transducer (brands of scanners and frequencies of transducers vary between studies)</p> <p><u>Participants' preparation:</u> Examination is performed in a dorsal lithotomy position with empty or half-full bladder; no bowel preparation is routinely required</p> <p><u>Protocol:</u> An ultrasound gel is applied to the tip of the transducer probe to create a lubricating, acoustically correct interface with the tissue. Scans are obtained by inserting the transducer (protected by disposable thin cover) into the vagina, followed by sequential movement of the probe within the vaginal canal to allow systematic evaluation of pelvic structures (uterus and adnexal regions; attention paid to the ovaries, pouch of Douglas, vesicouterine pouch and uterosacral ligament). The technique involves longitudinal,</p>	<p>TVS</p> <p>'transvaginal ultrasound'</p> <p>'transvaginal sonography'</p>
---	---	--

Table 2. Index tests - description and common abbreviations (Continued)

	transverse and angled movements of the probe with sliding up and down, back and forward to obtain both longitudinal and transversal scans of pelvic structures. Examination protocols vary between studies. Each examination is interpreted in real time and can be documented in printed photographs	
<ul style="list-style-type: none"> TVUS-BP (transvaginal ultrasonography with bowel preparation) 	Examination consists of TVUS combined with bowel preparation including the following: low-residue diet for 1-3 days, oral laxative on the eve of the examination, rectal enema within an hour before the examination or a combination of the above	
<ul style="list-style-type: none"> RWC-TVS (rectal water contrast transvaginal ultrasonography) 	<p>Examination consists of TVUS combined with bowel preparation and instillation of water contrast in rectum during TVUS; procedure does not require general anaesthesia</p> <p><u>Protocol:</u> After the transducer is introduced into the vagina, a flexible thin catheter (18-28 Ch) with a rubber balloon is inserted into the rectal lumen up to 20 cm from the anus (gel infused with lidocaine is used to facilitate passage of the catheter). Rectal water contrast of 100 to 300 mL of warm saline solution is instilled inside the balloon under ultrasonographic guidance to provide high-definition images of the rectal wall and its layers. Back flow of the solution is prevented by placement of a Klemmer forceps on the catheter. Images are obtained before, during and after saline injection</p>	<p>'transvaginal sonography with water-contrast in the rectum'</p> <p>'water-contrast in the rectum during transvaginal ultrasonography'</p>
<ul style="list-style-type: none"> SVG (sonovaginography) 	<p>Examination consists of TVUS combined with the introduction of saline solution or gel to the vagina to create an acoustical window between the transvaginal probe and surrounding structures and to distend the vaginal walls, permitting enhanced visualisation of pelvic structures</p> <p><u>Protocol:</u> Procedure involves introduction of a Foley catheter into the vagina followed by insertion of the transvaginal probe with further injection of 200-400 mL of saline through the catheter by the assistant. To prevent reflux of saline solution from the vagina, the vaginal canal is closed with the operator's hand. Alternative method involves placement of 20 mL of ultrasound gel into the posterior vaginal fornix with a plastic syringe, followed by insertion of a transvaginal probe. Reported procedure time ranges from 30 to 45 minutes</p>	'transvaginal sonography and acoustic window with intravaginal gel'
<ul style="list-style-type: none"> tg-TVUS (tenderness-guided TVUS) 	<p>Examination consists of TVUS combined with particular attention to the tender points evoked during examination</p> <p><u>Protocol:</u> Larger amount of ultrasound gel (~ 12 mL instead of the usual 4 mL) is introduced into the probe cover to create a stand-off for visualisation of the near-field area. The probe is inserted gently to avoid the risk of squeezing out the gel. After the initial sonographic evaluation, the participant is asked to inform the operator about the onset and site of any tenderness experienced during probe pressure within the posterior fornix. When tenderness is evoked, the sliding movement is stopped, and particular attention is paid to the painful site via gentle pressure with the probe's tip to detect endometriosis lesions. Reported procedure time is 15 to 20 minutes in cases of suspected lesions, but less time when the examination is negative</p>	
<ul style="list-style-type: none"> 3D-TVUS (3-dimensional transvaginal ultrasonography) 	<p><u>Equipment:</u> An ultrasound scanner equipped with 3D/4D imaging modes and a wide-band high resolution volume transvaginal transducer. The method enables the acquisition of ultrasonographic volumetric data that can be assessed off-line; in most institutions used as an adjunct to 2D US</p> <p><u>Protocol:</u> region-of-interest (ROI) is identified using a B-mode scan and a transvaginal volume transducer. During the volumetric scan, the transducer carries out a series of parallel scans of varying speeds focusing on the ROI. The anatomical ROI is visualised on the monitor as a graphic containing the 3 orthogonal planes. During volumetric scans, the investigator adopts some expedients such as positioning the probe near the anatomical ROI and</p>	

Table 2. Index tests - description and common abbreviations (Continued)

reducing or eliminating participant movements. The volume obtained is stored on a hard disk and displayed later using dedicated software

<ul style="list-style-type: none"> • Introital 3D-US (introital 3-dimensional ultrasound) 	<p>Examination is performed with the transducer placed on the perineum against the symphysis pubis (firmly but without causing significant discomfort). To acquire a correct volume, the symphysis pubis, urethra, vagina, and rectum should be visualised in the same image. Gain is adjusted and focal area is set to the region of interest, with the sweep angle set at 90 or 120 degrees to produce a multi-planar image in 3 planes: longitudinal, transverse and coronal</p>	
<p>TRUS (transrectal ultrasonography)</p>	<p>Equipment: An ultrasound scanner with a 2-dimensional axial and sagittal convex high-frequency probe with or without a rigid linear probe or a flexible endoscope with lateral view and a convex high frequency echo probe</p> <p>Participants' preparation: A low-residue diet for 3 days before the examination with or without laxatives and/or rectal enema is reported in some but not all studies; several groups described using general or local anaesthesia for the procedure, and some groups used no analgesia</p> <p>Protocol: A gel-filled rubber sheath or water-filled balloon is placed over the tip of the transducer to obtain better visibility. The transducer is inserted into the rectum and is advanced until the midline image of the cervix is visualised in the longitudinal view. Pelvic structures are evaluated by moving the transducer along its longitudinal axis and rotating it 130° to 140° along the main axis in both axial and longitudinal planes. Alternative technique includes insertion of the flexible probe into the sigmoid colon, over the aortic bifurcation and/or the upper part of the body of the uterus, with subsequent slow withdrawal, allowing optimum imaging of rectal and sigmoid colon walls/pelvic structures, with instillation of water into the intestinal lumen and alternating use of several frequencies (e.g. 5, 7.5, 12 MHz)</p>	<p>TRS (transrectal sonograph)</p> <p>Tr EUS (transrectal endoscopic ultrasonography)</p> <p>RES (rectal endoscopic sonography)</p> <p>REU (rectal endoscopic ultrasonography)</p>
<p>Other tests</p>		
<p>MDCT-e (multi-detector computerised tomography enema)</p>	<p>Equipment: multi-detector computed tomograph, which has a 2-dimensional array of detector elements that permits CT scanners to acquire multiple slices or sections simultaneously and greatly increase the speed of CT image acquisition (unlike the linear array of detector elements used in typical conventional and helical CT scanners)</p> <p>Participants' preparation: low-residue diet for 3 days and bowel preparation with an oral laxative day before the examination; intravenous injection of anti-peristaltic agent during the test</p> <p>Protocol: colonic distension performed by introducing about 2000 mL of water at 37°C into the left lateral decubitus position. All participants receive an intravenous injection of iodine-containing contrast. Participants are scanned in supine position from the dome of the diaphragm to the pubic symphysis in the portal phase (40 seconds after the arterial peak). Scan parameters (collimation, rotation time, tube voltage, effective mAs) differ between studies. Estimated radiation exposure is calculated by the scanner using CT dose index and is saved to the dose report. Both axial plane and multi-planar reconstructions (sagittal and coronal) are evaluated. Images are reviewed at a workstation</p>	<p>MSCTe (multi-slice computed tomography combined with colon distension by water enteroclysis)</p> <p>'Water enema CT'</p>
<p>¹⁸FDG-PET (fluorodeoxyglucose positron emission tomography)</p>	<p>Equipment: PET-computed tomograph</p> <p>Participants' preparation: Fasting for at least 6 hours before the test; ¹⁸FDG (a glucose analogue) injection 60 min before the test</p> <p>Protocol: Acquisition is performed with the participant in supine position, from mid-thigh to the base of the skull. No iodine-based contrast is administered. CT parameters reported in a single included study are 120 kV, 120 mA, pitch 1.5:1, speed 15 mm/rot. The PET element operates in 2D mode for 4 minutes per bed position. Attenuation correction is based on CT data</p>	

Table 2. Index tests - description and common abbreviations (Continued)

DCBE (double-contrast barium enema)	<u>Equipment</u> : motorised tilting radiographic table and standard equipment for fluoroscopic and radiological examination
	<u>Participants' preparation</u> : low-residue diet for 1-3 days before the examination with or without oral laxatives day before the procedure; an anti-peristaltic agent is administered intravenously at the outset of the examination to decrease bowel peristalsis
	<u>Protocol</u> : The procedure is performed in 2 steps to obtain double contrast and involves change of participant positions to ensure detailed visualisation of all intestinal segments. Barium sulphate contrast (600 to 800 mL) is instilled into rectum with a gravity pressure in the left lateral decubitus position. Once the barium reached the hepatic flexure, the colon was drained by gravity to remove as much barium as possible from the rectal ampulla without clearing completely the rectosigmoid colon of barium. Room air is then gently insufflated into the colon. Sequential views of the bowel are obtained. Each colonic segment is viewed in detail on spot radiographs and in magnification images. The procedure lasts 15 to 20 minutes

Table 3. Target conditions - types and anatomical distribution of endometriosis

Type of endometriosis	Description
Main clinical types of endometriosis	
Pelvic endometriosis	Endometriotic lesions, deep or superficial, located at any site in pelvic/abdominal cavity: on the peritoneum, fallopian tubes, ovaries, uterus, bowel, bladder or POD ^a
Ovarian endometriosis	Ovarian cysts lined by endometrial tissue (endometrioma)
DIE ^b	Deep endometriotic lesions extending more than 5 mm under the peritoneum located at any site of pelvic/abdominal cavity
Subtypes of deep endometriosis per anatomical localisation^c	
Posterior DIE	Deep endometriotic lesions involve ≥ 1 site of the posterior pelvic compartment (USL ^d RVSe, vaginal wall, bowel) and/or obliterate POD ^a
USL ^d endometriosis	Endometriotic lesions infiltrate uterosacral ligaments unilaterally or bilaterally
RVSe endometriosis	Deep endometriotic implants infiltrate the retroperitoneal area between posterior wall of vaginal mucosa and anterior wall of rectal muscularis
Vaginal endometriosis ^f	Endometriotic lesions infiltrate vaginal wall, particularly posterior vaginal fornix
POD ^a obliteration	Defined when the peritoneum of the POD ^a is only partially or no longer visible during surgery, and occurs as a result of adhesion formation; can be partial or complete, respectively
Bowel endometriosis	Endometriotic lesions infiltrating at least the muscular layer of the intestinal wall ileum - rectum; predominantly affects rectosigmoid colon
Rectosigmoid endometriosis	Endometriotic lesions infiltrating at least the muscular layer of the rectosigmoid colon; the most common form of bowel endometriosis
Anterior DIE	Deep endometriotic lesions located at any site of the anterior pelvic compartment (bladder \pm anterior pouch)

Table 3. Target conditions - types and anatomical distribution of endometriosis (Continued)

Rare types of endometriosis (not included in this review)

Bladder endometriosis	Endometriotic lesions infiltrating bladder muscularis propria
Ureteral endometriosis	Endometriotic lesions involving ureters
Extrapelvic/Atypical endometriosis	<p>Rare types of endometriosis involving various sites outside pelvic cavity, such as:</p> <p>CNS: cerebral endometriosis, extradural spinal endometriosis</p> <p>Thoracic: pleural endometriosis, pulmonary endometriosis, diaphragmatic endometriosis</p> <p>Abdominal: hepatic endometriosis, renal endometriosis, appendix endometriosis, pancreas endometriosis</p> <p>Musculoskeletal: abdominal wall endometriosis, umbilical endometriosis, pyramidalis muscle endometriosis, inguinal endometriosis, canal of Nuck endometriosis</p> <p>Perianal endometriosis, perineal endometriosis, extrapelvic endometriosis of sciatic nerve</p> <p>Subcutaneous endometriosis, operative scar endometriosis</p>

^aDIE: deep infiltrating endometriosis

^bPOD: pouch of Douglas

^cDefinitions of subtypes of DIE are adopted from Bazot 2007c. Additional definitions presented in the literature include 'Rectovaginal endometriosis (RVE)' defined as DIE that infiltrates the vagina, rectum and RVS and obliterates POD (Martin 2001) or 'deep retrocervical endometriosis' defined as involvement of USL, torus uterini, posterior vaginal fornix and/or RVS by endometriotic lesions (Abrao 2007).

^dUSL: uterosacral ligament

^eRVS: rectovaginal septum

^fVaginal endometriosis also defined as 'lesions infiltrating the anterior rectovaginal pouch, posterior vaginal fornix and retroperitoneal area between anterior rectovaginal pouch and posterior vaginal fornix (Chapron 2003a)

Table 4. Application of the QUADAS-2 tool for assessment of methodological quality of included studies
Domain 1 - Patient selection

Description	Describe methods of participant selection and characteristics of the included population
Type of bias assessed	Selection bias, spectrum bias
Review question	Women of reproductive age with clinically suspected endometriosis (symptoms, clinical examination ± presence of pelvic mass), scheduled for surgical exploration of pelvic/abdominal cavity for confirmation of the diagnosis ± treatment
Information collected	Study objectives, study population, selection (inclusion/exclusion criteria), study design, clinical presentation, age, number of enrolled and number available for analysis, setting, place and period of the study
Signalling question	Was a consecutive or random sample of participants enrolled?
Yes	If a consecutive sample or a random sample of eligible participants was included in the study

Table 4. Application of the QUADAS-2 tool for assessment of methodological quality of included studies (Continued)

No	If a non-consecutive sample or a non-random sample of eligible participants was included in the study
Unclear	All studies that did not specify enrolment as a consecutive or random sample of patients were classified as 'no'; therefore none of the included studies were classified as 'unclear'
Signalling question	Did the study avoid inappropriate exclusions?
Yes	If all participants with suspected endometriosis were included, with an exception for those not able to undergo an index test (e.g. virgins or genital tract anomalies for transvaginal imaging, claustrophobia for MRI) or unfit for surgery
No	If the study selected participants on the basis of particular clinical features (e.g. only suspected bowel involvement, were referred for treatment of deep endometriosis) or excluded participants with any co-morbidities, other than specified above
Unclear	If the study did not provide clear definition of selection (inclusion/exclusion) criteria and 'no' judgement was not applicable
Signalling question	Was a two-gate design avoided?
Yes	If the study had a single set of inclusion criteria, defined by the clinical presentation (i.e. only participants in whom the target condition is suspected) - a 'single-gate design'
No	If the study had more than 1 set of inclusion criteria with respect to clinical presentation (i.e. participants suspected of target condition, participants with alternative diagnosis in whom the target condition would not be suspected in clinical practice) - a 'two-gate' study design
Unclear	If it was unclear whether a 'two-gate design' was avoided
Risk of bias	Could the selection of participants have introduced bias?
High	If 'no' classification for any of the above 3 questions
Low	If 'yes' classification for 3 questions above
Unclear	If 'unclear' classification for any of the above questions and 'high risk' judgement were not applicable
Concerns about applicability	Are there concerns that included participants do not match the review question?
High	If the study population differed from the population defined in the review question in terms of demographic features and co-morbidity (e.g. studies with multiple sets of inclusion criteria with respect to clinical presentation, including healthy controls or alternative diagnosis controls that would not have undergone index test in real practice). We excluded studies in which participants were not in the reproductive age group, and most included studies were of 'single-gate' design; therefore, we expected few studies to be classified as 'high concern'
Low	If the study included only a clinically relevant population that would have undergone index test in real practice
Unclear	If this information was unclear
Domain 2 - Index test	
Description	Describe the index test, how it was conducted and interpreted

Table 4. Application of the QUADAS-2 tool for assessment of methodological quality of included studies (Continued)

Type of bias assessed	Test review bias, clinical review bias, interobserver variation bias
Review question	Any type of imaging modality
Information collected	Index test name, description of positive case definition by index test as reported, examiners (numbers, level of expertise, blinding), interobserver variability, conflicts of interest
Signalling question	<i>Were the index test results interpreted without knowledge of results of the reference standard?</i>
Yes	We excluded studies in which the index test was performed retrospectively after execution of the reference standard; therefore, all included studies were classified 'yes'
No	
Unclear	
Signalling question	<i>Did the study provide a clear prespecified definition of what was considered to be a 'positive result of the index test'?</i>
Yes	If study provided clear definition of positive findings, and this was defined before execution/interpretation of index test
No	If definition of the positive result was not provided, or if study described findings derived from the index test and not defined before its execution
Unclear	If it was unclear whether the criteria were prespecified
Signalling question	<i>Was the index test performed by a single operator or interpreted by consensus in a joint session?</i>
Yes	If test was performed/interpreted by single operator or was interpreted after collegial discussion of the case
No	If test was performed/interpreted by various operators for different participants
Unclear	If this information was unclear
Signalling question	<i>Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?</i>
Yes	If operators performing/interpreting the test were aware of suspected endometriosis and/or of the clinical history but were not aware of results of other imaging tests or of a previous diagnosis of endometriosis, including the results of previous surgeries
No	If operators performing/interpreting the test were informed of previously or recently surgically diagnosed endometriosis or were not blinded to results of other imaging tests or tests raising suspicion for endometriosis
Unclear	If this information was unclear
Risk of bias	<i>Could the conduct or interpretation of the index test have introduced bias?</i>
High	If 'no' classification for any of the above 4 questions
Low	If 'yes' classification for all the above 4 questions, or if 'unclear' classification for question 'Was the index test performed by a single operator or interpreted by consensus in a joint session?' and 'yes' classification for the remaining 3 questions

Table 4. Application of the QUADAS-2 tool for assessment of methodological quality of included studies (Continued)

Unclear	If 'unclear' classification at least for the question 'Did the study provide a clear pre-specified definition of what was considered to be a 'positive' result of index test?' or for the question 'Were the same clinical data available when the index test results were interpreted as would be available when the test is used in practice?' and 'high risk' judgement was not applicable
Concerns about applicability	Are there concerns that the index test, its conduct or its interpretation differs from the review question?
High	We did not consider studies in which index tests other than imaging modalities were included (or that excluded information on other index tests reported in addition to imaging modalities), or in which the index test looked at other target conditions not specified in the review (e.g. studies aimed at classifying pelvic masses as benign and malignant); therefore, none of the included studies was classified as 'high concern'
Low	We considered all types of imaging modalities as eligible; therefore, all included studies were classified as 'low concern', as anticipated
Unclear	Only studies with sufficient information on the index test were included; therefore, none of the included studies was classified as 'unclear concern'
Domain 3 - Reference standard	
Description	Describe the reference standard and how it was conducted and interpreted
Type of bias assessed	Verification bias, bias in estimation of diagnostic accuracy due to inadequate reference standard
Review question	Target condition - pelvic endometriosis, ovarian endometriosis, DIE overall or at specific anatomical sites; Reference standard - visualisation of endometriosis at surgery (laparoscopy or laparotomy) with or without histological confirmation
Information collected	Target condition, prevalence of target condition in the sample, reference standard, description of positive case definition by reference test as reported, examiners (numbers, level of expertise, blinding)
Signalling question	Is the reference standard likely to correctly classify the target condition?
Yes	If the study reported at least 1 of the following: surgical procedure described in sufficient detail and/or criteria for positive reference standard stated and/or the procedure was performed by the team with a high level of expertise in diagnosis/surgical treatment of the target condition
No	If the reference standard did not classify the target condition correctly; in the light of inclusion criteria and the nature of the reference standard, no studies were classified as 'no' for this item
Unclear	If information on execution of the reference standard or its interpretation or on operators was unclear
Signalling question	Were reference standard results interpreted without knowledge of results of the index tests?
Yes	If operators performing the reference test were unaware of the results of the index test
No	If operators performing the reference test were aware of the results of the index test
Unclear	If this information was unclear
Risk of bias	Could the reference standard, its conduct or its interpretation have introduced bias?
High	If 'no' classification for either of the above 2 questions

Table 4. Application of the QUADAS-2 tool for assessment of methodological quality of included studies (Continued)

Low	If 'yes' classification for both of the above 2 questions
Unclear	If 'unclear' classification for either of the above 2 questions and 'high risk' judgement was not applicable
Concerns about applicability	Are there concerns that the target condition as defined by the reference standard does not match the question?
High	We excluded studies in which participants did not undergo surgery for diagnosis of endometriosis; therefore, none of the included studies were classified as 'high concern'
Low	In the light of inclusion criteria, all studies were classified as 'low concern', as anticipated
Unclear	Only studies in which laparoscopy/laparotomy served as a reference test were included; therefore, no included studies were classified as 'unclear concern'
Domain 4 - Flow and timing	
Description	Describe any participants who did not receive the index tests or the reference standard, or who were excluded from the 2 × 2 table; describe the interval and any interventions between index tests and the reference standard
Type of bias assessed	Disease progression bias, bias of diagnostic performance due to missing data
Review question	Less than 12-month interval between index test and reference standard - endometriosis may progress over the time, so we had chosen an arbitrary time interval of 12 months as an acceptable time interval between the index test and surgical confirmation of the diagnosis
Information collected	Time interval between index test and reference standard, withdrawals (overall number reported and whether they were explained)
Signalling question	Was there an appropriate interval between index test and reference standard?
Yes	If time interval was reported and was less than 12 months
No	We excluded all studies for which the time interval was longer than 12 months; therefore, no included studies were classified as 'no' for this item
Unclear	If the time interval was not stated clearly but the study authors' description allowed one to assume that the interval was reasonably short
Signalling question	Did all participants receive the same reference standard?
Yes	In the light of inclusion criteria, all studies were classified as 'yes' for this item, as anticipated
No	
Unclear	
Signalling question	Were all participants included in the analysis?
Yes	If all participants were included in the analysis, or if participants were excluded because they did not meet inclusion criteria or if withdrawals were less than 5% of the enrolled population (arbitrary selected cut-off)
No	If any participants were excluded from the analysis because of uninterpretable results, because of inability to undergo index test or reference standard or for unclear reasons

Table 4. Application of the QUADAS-2 tool for assessment of methodological quality of included studies (Continued)

Unclear	No studies were classified as 'unclear' for this item
Risk of bias	Could the participant flow have introduced bias?
High	If 'no' classification for any of the above 3 questions
Low	If 'yes' classification for all of the above 3 questions
Unclear	If 'unclear' classification for any of the above 3 questions and 'high risk' judgement was not applicable

APPENDICES

Appendix 1. Electronic search strategies

Searches for the clinical studies

Database: MEDLINE (Ovid) <1946 to April, Week 2 2015 (20.04.2015)>

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R)

Search Strategy:

1. exp magnetic resonance imaging/ or exp ultrasonography/ or exp Imaging, Three-Dimensional/ or exp radiography/ (1114639)	Index test(s) set
2. ultraso\$.tw. or magnetic resonance imaging.tw. or MRI.tw. or imag\$.tw. (1020000)	
3. diagnos\$.tw. (1750239)	
4. or/1-3 (3048652)	
5. exp Endometriosis/ (17415)	Target condition set
6. Endometrio\$.tw. (21775)	
7. or/5-6 (25236)	
8. 4 and 7 (8107)	Combined sets
9. (animals not (humans and animals)).sh. (3931867)	
10. 8 not 9 (7391)	

Database: EMBASE (Embase.com) <1980 to 2015 April 20>

Search strategy:

1. Ecography/exp or radiodiagnosis/exp (1988601)	Index test(s) set
2. 'magnetic resonance imaging':ab,ti or MRI:ab,ti or imag*:ab,ti or ultraso*:de,ab,ti (1370683)	
3. diagnos*:ab,ti (2373625)	

Imaging modalities for the non-invasive diagnosis of endometriosis (Review)

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(Continued)

4. 'diagnostic accuracy':de or 'diagnostic test accuracy study':de or 'diagnostic value':de (298281)

5. or/1-4 (4437871)

6. Endometrio*:de,ab,ti (37439)

Target condition set

7. 'endometriosis'/exp/dm_di (4976)

8. or/6-7 (37439)

9. #5 and #8 (13500)

Combined sets

10. animal:de not (animal:de and human:de) (3861389)

11. #9 not #10 (12161)

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <April 2015 (20.4.2015)>

Search Strategy:

1. exp magnetic resonance imaging or exp ultrasonography or exp Imaging, Three-Dimensional or exp radiography (772)

Index test(s) set

2. (ultraso* or magnetic resonance imaging or MRI or imag*).tw. (36)

3. diagnos* (106503)

4. [mh diagnosis] (257329)

5. or/1-4 (310878)

6. exp endometriosis (142)

Target condition set

7. endometrio*.tw. (22)

8. [mh endometriosis] (553)

9. or/6-8 (681)

10. 5 and 9 (465)

Combined sets

11. (animals not (humans and animals)).sh. (36)

12. 10 not 11 (445)

Database CINAHL Plus with Full Text (EBSCOhost) <1980 to 20.04.2015>

Search strategy:

#	Query	Results	
S9	S3 AND S8	668	Combined sets
Search modes - Boolean/Phrase			

(Continued)

Search Screen - Advanced Search

S8	S4 OR S5 OR S6 OR S7	258011	Index test(s) set
S7	TX imag*	258011	
S6	TX ultraso*	58570	
S5	TX (magnetic resonance imaging or MRI)	58387	
S4	TX (biomarker* or marker*)	84857	
S3	S1 or S2	2841	Target condition set
S2	TX Endometrio*	2841	
S1	(MM "Endometriosis")	889	

Database: PsycINFO (Ovid) <1806 to April Week 2 2015 (20.04.2015)>

Search strategy:

1. endometriosis.tw. (174)

Database: Web of Science Core Collection (Thomson Reuters) <1900 to Present (20.04.2015)>

Search strategy:

1. Topic=(endometrio*) AND Topic=(diagnos* OR test* OR imag*); Timespan=All Years (7425)

Database: LILACS <20.04.2015>

Search strategy:

1. (tw:(endometriosis)) AND (tw:(diagnos*)) (420)

Database: OAlster (WorldCat.org) <20.04.2015>

Search strategy:

1. endometriosis and (marker* or biomarker*) (11)
2. endometriosis and diagnos* (446)

Database: TRIP <20.04.2015>

Search strategy:

1. (endometriosis and diagnos*) (1648)

Searches of trial registers for ongoing and registered trials
Database: ClinicalTrials.gov (US NIH) <20.04.2015>

Search strategy:

1. endometriosis (220)
2. endometriosis AND diagnosis (22)

Database: WHO International Clinical Trials Registry Platform (ICTRP) <20.04.2015>

Search strategy:

1. endometriosis (523)

Imaging modalities for the non-invasive diagnosis of endometriosis (Review)

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Searches for the reviews as source of references to potentially relevant studies

Database: MEDION <10.01.2014>

Search strategy:

ICP Code – female genital system (including breast), Signssymp – medical imaging, endoscopy and laparoscopy. Filter: systematic reviews of diagnostic studies (190)

Database: DARE (CRD) <20.04.2015>

Search strategy:

1. endometriosis (99)

PubMed, a 'Systematic Review' search under the 'Clinical Queries' link <20.04.2015>

Search strategy:

1. (endometriosis) AND systematic[sb] (418)

Category: Diagnosis; Scope: Broad

Searches for papers recently published and not yet indexed in the major databases

Search engine: PubMed <20.10.2014 to 20.04.2015>

Search strategy:

1. marker (14979)	Index test(s) set
2. test (61151)	
3. diagnos* (69743)	
4. biomarker (10806)	
5. or/1-4 (7943)	
Filters: Publication date from 2014/10/20 to 2015/04/20	
<hr/>	
6. Endometriosis (584)	Target condition set
Filters: Publication date from 2014/10/20 to 2015/04/20	
<hr/>	
7. 5 and 6 (267)	Combined sets
Filters: Publication date from 2014/10/20 to 2015/04/20	

Appendix 2. Direct comparisons of MRI methods for pelvic endometriosis

[Figure 9](#); [Figure 10](#); [Figure 11](#)

Appendix 3. Direct comparisons of imaging tests for ovarian endometriosis

[Figure 15](#); [Figure 16](#); [Figure 17](#)

Appendix 4. Direct comparisons of imaging tests for DIE/Posterior DIE

[Figure 23](#); [Figure 24](#); [Figure 25](#); [Figure 22](#)

Appendix 5. Direct comparisons of imaging tests for USL involvement by endometriosis

[Figure 29](#); [Figure 30](#); [Figure 31](#); [Figure 32](#)

Appendix 6. Direct comparisons of imaging tests for RVS involvement by endometriosis

[Figure 35](#); [Figure 36](#); [Figure 37](#); [Figure 38](#)

Appendix 7. Direct comparisons of imaging tests for Vaginal wall and POD involvement by endometriosis

[Figure 42](#); [Figure 43](#); [Figure 44](#); [Figure 45](#); [Figure 49](#)

Appendix 8. Direct comparisons of imaging tests for bowel involvement by endometriosis [rectosigmoid or overall bowel ileum - rectum]

[Figure 57](#); [Figure 58](#); [Figure 59](#); [Figure 60](#); [Figure 61](#); [Figure 62](#); [Figure 63](#); [Figure 64](#); [Figure 65](#); [Figure 66](#); [Figure 70](#); [Figure 71](#)

Appendix 9. List of abbreviations

Abbreviation	Description
2D	Two-dimensional
3D	Three-dimensional
CCDSO	Complete cul-de-sac obliteration
CDS	Cul-de-sac
CDSO	Cul-de-sac obliteration
CPP	Chronic pelvic pain
CSE	Conventional spin echo
CSE/TIFS	Conventional spin echo in combination with T1-weighted fat-suppressed
CSE/TIFS/Gd-TIFS	Conventional spin echo in combination with T1-weighted fat-suppressed and Gadolinium-enhanced TIFS
CT	Computed tomography
DCBE	Double-contrast barium enema
DE	Deep pelvic endometriosis
DIE	Deep infiltrating endometriosis or Deeply infiltrating endometriosis
DIPE	Deep infiltrating posterior endometriosis
DPE	Deep pelvic endometriosis
FDG PET CT	Fluorodeoxyglucose positron emission tomography/ computed tomography
FSE	Fast spin echo
Gd	Gadolinium
Gd-TIFS	Gadolinium-enhanced TIFS
GnRH	Gonadotropin-releasing hormone

(Continued)

MDCT-e	Multidetector computerized tomography enteroclysis
MRI	Magnetic resonance imaging
MRI-e	Magnetic resonance imaging enema
MSCT	Multi-slice computed tomography
MSCTe	Multi-slice computed tomography combined with colon distension by water enteroclysis
NPV	Negative predictive value
PACS	Picture archiving and communication system
PCDSO	Partial cul-de-sac obliteration
POD	Pouch of Douglas
PPV	Positive predictive value
RES	Rectal endoscopic sonography
RS	Rectosigmoid
RVS	Rectovaginal septum
RWC-TVSL	Rectal water contrast transvaginal ultrasonography
RWC-TVS	Rectal water contrast transvaginal ultrasonography
SVG	Sonovaginography
tg-TVUS	Tenderness-guided transvaginal ultrasound
TIFS	T1-weighted fat-suppressed
Tr EUS	Transrectal endoscopic ultrasonography
TRS	Transrectal sonography
TRUS	Transrectal ultrasonography
TVS	Transvaginal sonography
TVUS	Transvaginal ultrasonography
TVUS-BP	Transvaginal ultrasonography with bowel preparation
US	Ultrasound/Ultrasonography
USL	Uterosacral ligaments
USTV-PI	Transvaginal ultrasound with bowel preparation

CONTRIBUTIONS OF AUTHORS

Vicki Nisenblat and Louise Hull co-ordinated production of the protocol, performed the literature search, co-ordinated production of the review series and produced the first draft. Patrick Bossuyt provided advice on statistical methods for the review and performed the analyses. Cindy Farquhar critically reviewed the methodological aspects and participated in the study design. Neil Johnson contributed to the design of the review. All review authors contributed to revision and drafting of the review.

DECLARATIONS OF INTEREST

All review authors have declared no conflicts of interest.

SOURCES OF SUPPORT

Internal sources

- Cochrane Menstrual Disorders and Subfertility Group, University of Auckland, New Zealand.

Technical support

- The Robinson Institute, University of Adelaide, Other.

Access to academic resources

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

General scope: This review is a part of the review series arising from the same generic protocol. Sections were adjusted to the main topic of the review as follows.

- Background: The section on the index test was modified, and all information irrelevant to imaging tests removed. The 'Rationale' section was updated and now provides a clearer definition of triage diagnostic tests.
- Objectives.
 - * ☐ During revision of the literature on the topic, we identified a substantial body of studies looking at deep pelvic endometriosis as a separate entity, as well as at particular anatomical sites of deep pelvic endometriosis. We believe that assessment of different subtypes of endometriosis has clinical utility, which we have explained in the Background section under 'Rationale' and added to 'Objectives' as a secondary objective. We updated the definition of the target condition in the 'Methods' section, as mentioned below.
 - * We have updated the list of sources of heterogeneity.
- Methods.
 - * We updated criteria for considering studies for this review as follows.
 - ☐ Types of studies: We removed 'cohort' and 'case control' classifications and introduced the concept of 'single-gate design' and 'two-gate' design'. This was defined as the presence of a single or multiple set of inclusion criteria with regard to the clinical condition or the reference standard. We found this classification more informative for describing diagnostic studies, all of which are cross-sectional in nature. We limited inclusion criteria to studies with a single set of inclusion criteria by reference standard (i.e. all women who underwent abdominal surgery) but included single or multiple sets of inclusion criteria by clinical presentation (i.e. women with suspected endometriosis or other indications for abdominal surgery), referring to these as 'single-gate' and 'two-gate' designs, respectively.
 - ☐ We modified index tests to pertain only to imaging modalities, and we updated the table listing tests of interest ([Table 2](#)) accordingly.
 - ☐ Target conditions also included deep pelvic endometriosis in view of the growing body of literature on this condition as a separate entity and its diagnostic importance in optimising the surgical approach. Target conditions per different pelvic compartments and anatomical sites are presented in [Table 3](#).
 - ☐ Spectrum of disease: Following ad hoc observation, we included studies that involved only a selected population of women with endometriosis (i.e. specific rASRM stages) in view of emerging evidence on the poor correlation of this classification with infertility and pain symptoms. Exclusion of such studies could result in loss of potentially important diagnostic information from otherwise eligible publications. When possible, we aimed to address the impact of including these studies in investigations of heterogeneity.
 - * Search methods for identification of studies.
 - ☐ In the protocol, we stated that we would identify the grey literature (unpublished studies including conference proceedings and reports), and we defined specific search strategies. In practice, the paucity of relevant data that was available from abstracts made it impossible for us to apply selection criteria and methodological quality judgement to these studies. Identification of this type of study and attempts to obtain necessary information directly from study investigators were anticipated to increase the already

labour intense work involved in preparation of this review. Therefore, by consensus between key review authors, we removed already identified unpublished studies and did not complete an intended search for unpublished material.

☐ We updated search strings for imaging tests by applying the same principles as presented in the protocol ([Appendix 1](#)).

* Assessment of methodological quality: We tailored the QUADAS-2 tool for the topic of the review. We outlined differences between the original QUADAS-2 tool and the tool designed for this review in the relevant section under Methods.

* Analysis.

☐ We amended the section on statistical methods and tailored it to the types of tests included in the review.

☐ We performed no sensitivity analyses and no assessment of heterogeneity because data for most tests were insufficient, except for TVUS for DIE, RVS and rectosigmoid endometriosis.

☐ When we judged test performance against predetermined diagnostic criteria, we considered the point estimates of sensitivity and specificity as most informative of test performance. We acknowledge that tests with point estimates that did not reach predetermined criteria with confidence intervals (CIs) that contained values above the threshold could have diagnostic value. Furthermore, tests with point estimates that reached the criteria but with CIs that contained values below the threshold could have overestimated diagnostic value. If the range of CIs rather than the point estimates of data were used, the predetermined cut-off would become meaningless. Therefore, we did not consider CIs when qualifying test performance but utilised this information when interpreting the reliability of obtained data.

• We changed the list and order of review authors to accurately reflect author contributions to the review.

NOTES

A single review on non-invasive tests for diagnosis of endometriosis was planned but was split into several smaller reviews to facilitate data handling and interpretation as a result of the abundance and diversity of suggested tests. We generated this review from a generic protocol, which had been designed for all reviews in these series. Other reviews from this series include (1) Endometrial biomarkers for the non-invasive diagnosis of endometriosis; (2) Urinary biomarkers for the non-invasive diagnosis of endometriosis; (3) Blood biomarkers for the non-invasive diagnosis of endometriosis; and (4) Combined tests for the non-invasive diagnosis of endometriosis, the last of which is a summarising review of the series.

INDEX TERMS

Medical Subject Headings (MeSH)

Chronic Disease; Cross-Sectional Studies; Diagnostic Imaging [*methods]; Endometriosis [*diagnosis] [pathology]; Magnetic Resonance Imaging; Ovarian Diseases [diagnosis] [surgery]; Pelvis; Positron-Emission Tomography; Randomized Controlled Trials as Topic; Sensitivity and Specificity; Ultrasonography

MeSH check words

Female; Humans